



```
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 11823-37-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other nucleic acid
; DESCRIPTION: synthetic DNA
US-08-458-516-21
Alignment Scores:
Pred. No.: 4.9 Length: 120
Score: 49.00 Matches: 7
Percent Similarity: 71.43% Conservative: 3
Best Local Similarity: 50.00% Mismatches: 4
Query Match: 33.33% Indels: 0
DB: 1 Gaps: 0
US-10-014-101B-32 (1-28) x US-08-458-516-21 (1-120)
Qy 9 TyrProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22
Db 75 TACCCTGTGGCAAAATAAGTCGCGAAATCATCTGGCTGCA 116
RESULT 33
US-09-270-767-1938
; Sequence 1938, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1938
; LENGTH: 350
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-1938
Alignment Scores:
Pred. No.: 19.8 Length: 350
Score: 49.00 Matches: 9
Percent Similarity: 62.50% Conservative: 6
Best Local Similarity: 37.50% Mismatches: 9
Query Match: 33.33% Indels: 0
DB: 4 Gaps: 0
US-10-014-101B-32 (1-28) x US-09-270-767-1938 (1-350)
Qy 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20
Db 230 AGCGTCAGCCAGCTTGGCAGATTGTGTTATTGAGCGGTCCAGTTGGGATCAGCGACTC 289
RESULT 34
US-09-270-767-17220
; Sequence 17220, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
```

```
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17220
; LENGTH: 350
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-17220
Alignment Scores:
Pred. No.: 19.8 Length: 350
Score: 49.00 Matches: 9
Percent Similarity: 62.50% Conservative: 6
Best Local Similarity: 37.50% Mismatches: 9
Query Match: 33.33% Indels: 0
DB: 4 Gaps: 0
US-10-014-101B-32 (1-28) x US-09-270-767-17220 (1-350)
Qy 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20
Db 230 AGCGTCAGCCAGCTTGGCAGATTGTGTTATTGAGCGGTCCAGTTGGGATCAGCGACTC 289
RESULT 35
US-09-543-681A-925/c
; Sequence 925, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 925
; LENGTH: 858
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-925
Alignment Scores:
Pred. No.: 64 Length: 858
Score: 49.00 Matches: 10
Percent Similarity: 71.43% Conservative: 5
Best Local Similarity: 47.62% Mismatches: 6
Query Match: 33.33% Indels: 0
DB: 4 Gaps: 0
US-10-014-101B-32 (1-28) x US-09-543-681A-925 (1-858)
Qy 6 AlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSerAlaMetIlePro 25
Db 794 GCGATTCTTGGCAGCCCAATAGATTAAAGCCAAAATAGAAATAGTCTGTCTACCT 735
Qy 26 Glu 26
Db 734 GAG 732
RESULT 36
US-08-728-603-18
; Sequence 18, Application US/08728603
; Patent No. 6093806
; GENERAL INFORMATION:
; APPLICANT: Cesarman, Ethel
```

```
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 813
; LENGTH: 6439
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-813

Alignment Scores:
Pred. No.: 731 Length: 6439
Score: 49.50 Matches: 9
Percent Similarity: 69.23% Conservative: 9
Best Local Similarity: 34.62% Mismatches: 7
Query Match: 33.67% Indels: 1
DB: 4 Gaps: 1

US-10-014-101B-32 (1-28) x US-09-902-540-813 (1-6439)

QY 3 SerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsn---ArgMetSer 21
Db 4083 GCTGGTATGCCAGCAGAGATTCCCTCAACCCCAATGGATGGAGCAACGCCCAACTGCAC 4142
QY 22 AlaMetIleProGluIle 27
Db 4143 GCCTCATGAGGATGTC 4160

RESULT 30
US-09-949-016-13465/c
; Sequence 13465, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13465
; LENGTH: 27579
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13465

Alignment Scores:
Pred. No.: 4.89e+03 Length: 14115
Score: 49.50 Matches: 10
Percent Similarity: 59.09% Conservative: 3
Best Local Similarity: 45.45% Mismatches: 2
Query Match: 33.67% Indels: 7
DB: 4 Gaps: 1

US-10-014-101B-32 (1-28) x US-09-949-016-17490 (1-14115)

QY 4 GlyLeuAlaLeuLeuTyProThrAsnArg-----AsnLysTrp 16
Db 14577 GCCTGACTCTTTTATTCCTTCAACAGATGAGCTGACATGGGCTATCAACGGGTGG 14636
QY 17 AspAsn 18
Db 14637 GAAAC 14642

RESULT 32
US-08-458-516-21
; Sequence 21, Application US/08458516
; Patent No. 5777085
; GENERAL INFORMATION:
; APPLICANT: Co, Man Sung
; APPLICANT: Tso, J. Yun
; TITLE OF INVENTION: Humanized Antibodies Reactive with
; TITLE OF INVENTION: GPIIB/IIIA
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: William M. Smith
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,516
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: US 08/059,159
; FILING DATE: 03-MAY-1993
; ATTORNEY/AGENT INFORMATION:
```

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/149,718  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/660,487  
FILING DATE:  
APPLICATION NUMBER: 08/480,653  
FILING DATE: June 7, 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Fabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: ANS101CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (404)-873-8794  
TELEFAX: (404)-873-8795  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2085 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1-2085  
OTHER INFORMATION: /function= "coding region for APP695."

US-09-149-718-1

Alignment Scores:

Pred. No.: 138 Length: 2085

Score: 50.00 Matches: 11

Percent Similarity: 55.56% Conservative: 4

Best Local Similarity: 40.74% Mismatches: 12

Query Match: 34.01% Indels: 0

DB: 4 Gaps: 0

US-10-014-101b-32 (1-28) x US-09-149-718-1 (1-2085)

Qy 2 AlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSer 21

Db 1643 AGTTGGGGGCTGACTGTGCGCCACACAGAAACGAAGTTGAGCCTGTTGATGCC 1702

Qy 22 AlaMetileProGlulileAsp 28

Db 1703 GCCCTGCTGCCGACCGGAGAC 1723

RESULT 27

US-09-902-540-611

; Sequence 611, Application US/09902540

; Patent No. 6833447

; GENERAL INFORMATION:

; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Wiegand, Roger C.

; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

; FILE REFERENCE: 38-10(15849)B

; CURRENT APPLICATION NUMBER: US/09/902,540

; CURRENT FILING DATE: 2001-07-10

; PRIOR APPLICATION NUMBER: 60/217,883

; PRIOR FILING DATE: 2000-07-10

; NUMBER OF SEQ ID NOS: 16825

; SEQ ID NO 611

; LENGTH: 2388

; TYPE: DNA

; ORGANISM: Myxococcus xanthus

US-09-902-540-611

Alignment Scores:

Pred. No.: 164 Length: 2388

Score: 50.00 Matches: 11

Percent Similarity: 56.00% Conservative: 3

Best Local Similarity: 44.00% Mismatches: 11

Query Match: 34.01% Indels: 0

DB: 4 Gaps: 0

US-10-014-101b-32 (1-28) x US-09-902-540-611 (1-2388)

Qy 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20

Db 382 TCCAACCTCCGGCTTTCGGCTGTCGTACCCAGCCCTCCAGTTCCTGGGCGCCGCC 441

Qy 21 SerAlaMetilePro 25

Db 442 TCCGCTCGGAGACCC 456

RESULT 28

US-09-949-016-16923

; Sequence 16923, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 16923

; LENGTH: 390416

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-16923

Alignment Scores:

Pred. No.: 1,248+05 Length: 390416

Score: 50.00 Matches: 8

Percent Similarity: 73.33% Conservative: 3

Best Local Similarity: 53.33% Mismatches: 4

Query Match: 34.01% Indels: 0

DB: 4 Gaps: 0

US-10-014-101b-32 (1-28) x US-09-949-016-16923 (1-390416)

Qy 13 ArgAsnLysTrpAspAsnArgMetSerAlaMetileProGlulile 27

Db 228116 AAAACAATGCGAGGAAAGTTCAGCTGTATTTCAGGAATA 228160

RESULT 29

US-09-902-540-813

; Sequence 813, Application US/09902540

; Patent No. 6833447

; GENERAL INFORMATION:

; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Wiegand, Roger C.

; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

; FILE REFERENCE: 38-10(15849)B

; CURRENT APPLICATION NUMBER: US/09/902,540

; CURRENT FILING DATE: 2001-07-10

; PRIOR APPLICATION NUMBER: 60/217,883

```
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 4
; OTHER INFORMATION: seq FSPMLLGMGGCLP/GF
; NAME/KEY: polyA site
; LOCATION: 536..547
US-09-663-600A-137

Alignment Scores:
Pred. No.: 24 Length: 547
Score: 50.00 Matches: 12
Percent Similarity: 60.00% Conservative: 3
Best Local Similarity: 48.00% Mismatches: 6
Query Match: 34.01% Indels: 4
DB: 4 Gaps: 1

US-10-014-101B-32 (1-28) x US-09-663-600A-137 (1-547)
QY 1 SerAlaSerGlyLeuAlaLeuLeuTyrProThrAsnArgAsnLysTrpAspAsnArgMet 20
Db 400 GCAGGCTCTGGGCTTTTCTCTTTATGTACT-----TGGAAATGGCGGGTG 447
QY 21 SerAlaMetIlePro 25
Db 448 CCTGCCTGGATTCT 462

RESULT 24
US-09-663-600A-43
; Sequence 43, Application US/09663600A
; Patent No. 6573068
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS
; FILE REFERENCE: 31.US3.CIP
; CURRENT APPLICATION NUMBER: US/09/663,600A
; CURRENT FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/191,997
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/099,273
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: Patent.pm
; SEQ ID NO 43
; LENGTH: 648
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: sig peptide
; LOCATION: 460..555
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 4
; OTHER INFORMATION: seq FSPMLLGMGGCLP/GF
; NAME/KEY: polyA signal
; LOCATION: 614..619
; NAME/KEY: polyA site
; LOCATION: 635..648
US-09-663-600A-43

Alignment Scores:
Pred. No.: 29.9 Length: 648
Score: 50.00 Matches: 12
Percent Similarity: 60.00% Conservative: 3

; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 4
; OTHER INFORMATION: seq FSPMLLGMGGCLP/GF
; NAME/KEY: polyA site
; LOCATION: 536..547
US-09-663-600A-137

Alignment Scores:
Pred. No.: 24 Length: 547
Score: 50.00 Matches: 12
Percent Similarity: 60.00% Conservative: 3
Best Local Similarity: 48.00% Mismatches: 6
Query Match: 34.01% Indels: 4
DB: 4 Gaps: 1

US-10-014-101B-32 (1-28) x US-09-663-600A-137 (1-547)
QY 1 SerAlaSerGlyLeuAlaLeuLeuTyrProThrAsnArgAsnLysTrpAspAsnArgMet 20
Db 400 GCAGGCTCTGGGCTTTTCTCTTTATGTACT-----TGGAAATGGCGGGTG 447
QY 21 SerAlaMetIlePro 25
Db 448 CCTGCCTGGATTCT 462

RESULT 25
US-09-248-796A-738
; Sequence 738, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 738
; LENGTH: 1008
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-738

Alignment Scores:
Pred. No.: 53.2 Length: 1008
Score: 50.00 Matches: 12
Percent Similarity: 60.71% Conservative: 5
Best Local Similarity: 42.86% Mismatches: 2
Query Match: 34.01% Indels: 9
DB: 4 Gaps: 1

US-10-014-101B-32 (1-28) x US-09-248-796A-738 (1-1008)
QY 3 SerGlyLeuAlaLeuLeuTyrProThrAsnArgAsnLys-----15
Db 559 AGTGGTATGCTATGCTATACCAAGGATCGAAAAGATATTTTACAAACATTTTGG 618
QY 16 ----TrpAspAsnArgMetSer 21
Db 619 AATTTGGGATTCCTCGGTGAGT 640

RESULT 26
US-09-149-718-1
; Sequence 1, Application US/09149718
; Patent No. 6717031
; GENERAL INFORMATION:
; APPLICANT: Dora K. Games, Dale B. Schenk, Lisa C. McConlogue,
; APPLICANT: Peter A. Seubert, and Russell E. Rydel
; TITLE OF INVENTION: Method For Identifying Alzheimer's Disease
; TITLE OF INVENTION: Therapeutics Using Transgenic Animal Models
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabet
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: GA
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
```

```
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15820
; LENGTH: 53394
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(53394)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15820

Alignment Scores:
Pred. No.: 6,41e+03 Length: 53394
Score: 51.00 Matches: 11
Percent Similarity: 65.22% Conservative: 4
Best Local Similarity: 47.83% Mismatches: 8
Query Match: 34.69% Indels: 0
DB: 4 Gaps: 0

US-10-014-101B-32 (1-28) x US-09-949-016-15820 (1-53394)
QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20
Db 41543 AGTGCTCTGGAGTGGCCGATCTGATGTTCCACGCGTCCCTCTGGATCGCGGATG 41484
QY 21 SerAlaMet 23
Db 41483 GACGCGCTC 41475

RESULT 21
US-09-596-002-32/c
; Sequence 32, Application US/09596002
; Patent No. 6632636
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert, E.
; APPLICANT: Patterson, Chandra
; APPLICANT: Berg, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: PM-0008-4 US
; CURRENT APPLICATION NUMBER: US/09/596,002
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 60/140,121
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 32
; LENGTH: 62909
; TYPE: DNA
; ORGANISM: M. catarrhalis
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte template ID No. 6632636 32
; PUBLICATION INFORMATION:
US-09-596-002-32

Alignment Scores:
Pred. No.: 7,94e+03 Length: 62909
Score: 51.00 Matches: 8
Percent Similarity: 76.47% Conservative: 5
Best Local Similarity: 47.06% Mismatches: 4
Query Match: 34.69% Indels: 0
DB: 4 Gaps: 0

US-10-014-101B-32 (1-28) x US-09-596-002-32 (1-62909)
QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAsp 17
Db 6032 TCTACACAGGCATATCCATAAGTATCTCTTAAATATCAATGAATGGGAT 5982

RESULT 22
US-09-902-540-6994/c

; Sequence 6994, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 6994
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-6994

Alignment Scores:
Pred. No.: 21.9 Length: 510
Score: 50.00 Matches: 11
Percent Similarity: 56.00% Conservative: 3
Best Local Similarity: 44.00% Mismatches: 11
Query Match: 34.01% Indels: 0
DB: 4 Gaps: 0

US-10-014-101B-32 (1-28) x US-09-902-540-6994 (1-510)
QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20
Db 131 TCCAACTCCGGCTTTCGGCTGTCGTACCCAGCGCTCCAGTCTCTGGCGCCGCCGCC 72
QY 21 SerAlaMetIlepro 25
Db 71 TCCGCTCGGAGACCC 57

RESULT 23
US-09-663-600A-137
; Sequence 137, Application US/09663600A
; Patent No. 6573068
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS
; FILE REFERENCE: 31.US3.CIP
; CURRENT APPLICATION NUMBER: US/09/663,600A
; CURRENT FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/191,997
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/099,273
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: Patent.pm
; SEQ ID NO 137
; LENGTH: 547
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 359..454
```

```

; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15817
; LENGTH: 53394
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(53394)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15817

Alignment Scores:
Pred. No.: 6.41e+03 Length: 53394
Score: 51.00 Matches: 11
Percent Similarity: 65.22% Conservative: 4
Best Local Similarity: 47.83% Mismatches: 8
Query Match: 34.69% Indels: 0
Gaps: 0

US-10-014-101b-32 (1-28) x US-09-949-016-15817 (1-53394)
QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20
Db 41543 AGTGCCTCTGGAGTGGCCGATCTGATGTTGTCACGCGTCCCTCTCGGAATCGGCGGATG 41484
QY 21 SerAlaMet 23
Db 41483 GACGCGCTC 41475

RESULT 19
US-09-949-016-15819/c
; Sequence 15819, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15819
; LENGTH: 53394
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(53394)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15819

Alignment Scores:
Pred. No.: 6.41e+03 Length: 53394
Score: 51.00 Matches: 11
Percent Similarity: 65.22% Conservative: 4
Best Local Similarity: 47.83% Mismatches: 8
Query Match: 34.69% Indels: 0
Gaps: 0

US-10-014-101b-32 (1-28) x US-09-949-016-15819 (1-53394)
QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20
Db 41543 AGTGCCTCTGGAGTGGCCGATCTGATGTTGTCACGCGTCCCTCTCGGAATCGGCGGATG 41484
QY 21 SerAlaMet 23
Db 41483 GACGCGCTC 41475

RESULT 18
US-09-949-016-15818/c
; Sequence 15818, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15818
; LENGTH: 53394
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(53394)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15818

Alignment Scores:
Pred. No.: 6.41e+03 Length: 53394
Score: 51.00 Matches: 11
Percent Similarity: 65.22% Conservative: 4
Best Local Similarity: 47.83% Mismatches: 8
Query Match: 34.69% Indels: 0
Gaps: 0

US-10-014-101b-32 (1-28) x US-09-949-016-15818 (1-53394)
QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20
Db 41543 AGTGCCTCTGGAGTGGCCGATCTGATGTTGTCACGCGTCCCTCTCGGAATCGGCGGATG 41484
QY 21 SerAlaMet 23
Db 41483 GACGCGCTC 41475

RESULT 20
US-09-949-016-15820/c
; Sequence 15820, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15818
; LENGTH: 53394
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(53394)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15820
```

```
STATE: Iowa
COUNTRY: USA
ZIP: 50309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/123,761A
  FILING DATE: 17-SEP-1993
  CLASSIFICATION: 800
  ATTORNEY/AGENT INFORMATION:
    NAME: Roth, Michael J.
    REGISTRATION NUMBER: 29,342
    REFERENCE/DOCKET NUMBER: 212-US
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (515) 248-4800
    TELEFAX: (515) 248-4844
  INFORMATION FOR SEQ ID NO: 1:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 5198 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: double
      TOPOLOGY: linear
    MOLECULE TYPE: DNA (genomic)
    US-08-123-761A-1

Alignment Scores:
Pred. No.: 206 Length: 5198
Score: 52.00 Matches: 11
Percent Similarity: 66.67% Conservative: 5
Best Local Similarity: 45.83% Mismatches: 8
Query Match: 35.37% Indels: 0
DB: 1 Gaps: 0

US-10-014-101B-32 (1-28) x US-08-123-761A-1 (1-5198)
QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20
Db 4527 ACCGCTCCGCTCTGACGAGATCAATCCAACTCAACGAGAAACGTAATCGCAG 4468
QY 21 SerAlaMetIle 24
Db 4467 TCAGTTCTTG 4456

RESULT 15
US-09-949-016-17547/c
; Sequence 17547, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17547
; LENGTH: 84870
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(84870)
; OTHER INFORMATION: n = A,T,C or G

US-09-949-016-17547
Alignment Scores:
Pred. No.: 7,91e+03 Length: 84870
Score: 52.00 Matches: 11
Percent Similarity: 60.71% Conservative: 6
Best Local Similarity: 39.29% Mismatches: 11
Query Match: 35.37% Indels: 0
DB: 4 Gaps: 0

US-10-014-101B-32 (1-28) x US-09-949-016-17547 (1-84870)
QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20
Db 4466 TCATTACTAACTGGCCCTATTACATGCTCAACTAATAACAATGGGAATAGAAGC 4407
QY 21 SerAlaMetIleProGluIleAsp 28
Db 4406 ACTAATTTCTTACCTTTTCATAGAC 4383

RESULT 16
US-09-657-252-3/c
; Sequence 3, Application US/09657252
; Patent No. 6500643
; GENERAL INFORMATION:
; APPLICANT: Wu, Dong-Hai
; APPLICANT: Gu, Yunrong
; APPLICANT: Millard, William
; APPLICANT: He, Yun-Je
; TITLE OF INVENTION: Human High Affinity Choline Transporter cDNA
; FILE REFERENCE: MEB00-639
; CURRENT APPLICATION NUMBER: US/09/657,252
; CURRENT FILING DATE: 2000-09-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 1893
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1893)
; US-09-657-252-3

Alignment Scores:
Pred. No.: 81.7 Length: 1893
Score: 51.00 Matches: 11
Percent Similarity: 65.22% Conservative: 4
Best Local Similarity: 47.83% Mismatches: 8
Query Match: 34.69% Indels: 0
DB: 4 Gaps: 0

US-10-014-101B-32 (1-28) x US-09-657-252-3 (1-1893)
QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20
Db 1379 AGTGCCTCTGGAGTGGCGGATCTGATGTTGCCAGGCTCCTCTCGAATCGGCGATG 1320
QY 21 SerAlaMet 23
Db 1319 GACGCGCTC 1311

RESULT 17
US-09-949-016-15817/c
; Sequence 15817, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
```

```
; Sequence 13046, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: GL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13046
; LENGTH: 28555
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(28555)
; OTHER INFORMATION: n = A, T, C or G
US-09-949-016-13046

Alignment Scores:
Pred. No.: 866 Length: 28555
Score: 54.00 Matches: 11
Percent Similarity: 56.52% Conservative: 2
Best Local Similarity: 47.83% Mismatches: 2
Query Match: 36.73% Indels: 8
DB: 4 Gaps: 1

US-10-014-101B-32 (1-28) x US-09-949-016-13046 (1-28555)

QY 3 SerGlyLeuAlaLeuLeu-----TyrProThrAsnArgAsn 14
|||:::|||||
Db 27369 TCGGGATGATTTGTTGCTTCAGAACCTTCCTCCATTACCTGTGATATAAAC 27310
|||:::|||||

QY 15 LysTrpAsp 17
|||:::|||||
Db 27309 AAATGGGAC 27301

RESULT 13
US-08-261-206A-71/c
; Sequence 71, Application US/08261206A
; Patent No. 5574007
; GENERAL INFORMATION:
; APPLICANT: Zushi, Mitichitaka
; APPLICANT: Gomi, Komakazu
; APPLICANT: Yamamoto, Shuji
; APPLICANT: Suzuki, Koji
; APPLICANT: Matsuda, Akio
; TITLE OF INVENTION: A Polypeptide Capable of Interacting
; WITH THROMBIN
; NUMBER OF INVENTION: with Thrombin
; NUMBER OF SEQUENCES: 80
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 301 N. Washington St.
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22046-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/261,206A
; FILING DATE:
```

```
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/740,492
; FILING DATE: 03-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Svensson, Leonard R.
; REGISTRATION NUMBER: 30330
; REFERENCE/DOCKET NUMBER: 216-275P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-241-1300
; TELEFAX: 703-241-2848
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 71:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3306 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Acremonium chrysogenum
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..3306
; OTHER INFORMATION: /label= PGK gene
; OTHER INFORMATION: /note= "Nucleotide sequence of region A in Figure
; 59. The sequence is presented as Figure 61."
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1252..1317
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1463..1883
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1948..2715
; FEATURE:
; NAME/KEY: CDS
; LOCATION: Join(1252..1317, 1463..1883, 1948..2714)
US-08-261-206A-71

Alignment Scores:
Pred. No.: 63.1 Length: 3306
Score: 53.50 Matches: 13
Percent Similarity: 53.12% Conservative: 4
Best Local Similarity: 40.62% Mismatches: 8
Query Match: 36.39% Indels: 7
DB: 1 Gaps: 1

US-10-014-101B-32 (1-28) x US-08-261-206A-71 (1-3306)

QY 1 SerAlaSerGly-----LeuAlaLeuTyrProThrAsnArg 13
|||:::|||||
Db 1653 TCGAGCTCGGAGCAGCGGCTGACGAGTACTTGGGGTGGGGGTGCGGCGG 1594
|||:::|||||

QY 14 AsnLysTrpAspAsnArgMetSerAlaMetIlePro 25
|||:::|||||
Db 1593 CCAAGGTGGGACATGAGGATGACGCGCTTGGCGCCA 1558
|||:::|||||

RESULT 14
US-08-123-761A-1/c
; Sequence 1, Application US/08123761A
; Patent No. 5589611
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
; APPLICANT: Johal, Gurmukh S.
; TITLE OF INVENTION: A DISEASE RESISTANCE GENE FROM MAIZE AS
; A SELECTABLE MARKER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
; STREET: 700 Capital Square, 400 Locust Street
; CITY: Des Moines
```

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 413-3000  
TELEFAX: (703) 413-2220  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2763 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE: Lactococcus cremoris  
STRAIN: IL964  
US-08-248-466B-2

Alignment Scores:  
Pred. No.: 27.6 Length: 2763  
Score: 55.00 Matches: 12  
Percent Similarity: 58.33% Conservative: 2  
Best Local Similarity: 50.00% Mismatches: 10  
Query Match: 37.41% Indels: 0  
DB: 1 Gaps: 0

US-10-014-101B-32 (1-28) x US-08-248-466B-2 (1-2763)

QY 5 LeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSerAlaMetIle 24  
Db 540 TTAGCCCTTAATTCCTTAACCTGCTCATRAAGTGGATAACGTTTTCGTGAAGCGTG 481

QY- 25 ProGluIleAsp 28  
Db 480 CCAATTATTGAC 469

## RESULT 10

US-09-949-016-13407/c  
Sequence 13407, Application US/09949016  
Patent No. 6812339

GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 13407

LENGTH: 32616

TYPE: DNA

ORGANISM: Human

US-09-949-016-13407

Alignment Scores:  
Pred. No.: 694 Length: 32616  
Score: 55.00 Matches: 11  
Percent Similarity: 66.67% Conservative: 3  
Best Local Similarity: 52.38% Mismatches: 7  
Query Match: 37.41% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-32 (1-28) x US-09-949-016-13407 (1-32616)

QY 1 SerAlaSerGlyLeuAlaLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20  
Db 17874 TCACCTCAGGCGCTAGCAATCTCTGGATGACACCTCTGAGAGCAATGGACAGAGGGAA 17815

QY 21 Ser 21  
Db 17814 AGT 17812

## RESULT 11

US-09-023-655-543  
Sequence 543, Application US/09023655  
Patent No. 6607879

GENERAL INFORMATION:  
APPLICANT: Cocks, Benjamin G.

APPLICANT: Susan G. Stuart

APPLICANT: Jeffrey J. Seilhamer

TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE

EXPRESSION

NUMBER OF SEQUENCES: 1508

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/023,655

FILING DATE: HEREWITH

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37,071

REFERENCE/DOCKET NUMBER: PA-0001 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 543:

SEQUENCE CHARACTERISTICS:

LENGTH: 1031 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: STOMNOT01

CLONE: 221877

US-09-023-655-543

Alignment Scores:

Pred. No.: 11.3 Length: 1031

Score: 54.00 Matches: 11

Percent Similarity: 56.52% Conservative: 2

Best Local Similarity: 47.83% Mismatches: 2

Query Match: 36.73% Indels: 8

DB: 4 Gaps: 1

US-10-014-101B-32 (1-28) x US-09-023-655-543 (1-1031)

QY 3 SerGlyLeuAlaLeuLeu-

Db 281 TCTGGATGCATTGTTGTCCTTCAGACCCCTTCCTCCCATTTACCTGTCATATAAAC 340

QY 15 LysTrpAsp 17

Db 341 AAATGGGAC 349

## RESULT 12

US-09-949-016-13046/c

```
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13546
; LENGTH: 163664
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(163664)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-13546

Alignment Scores:
Pred. No.: 4.68e+03 Length: 163664
Score: 55.50 Matches: 12
Percent Similarity: 63.64% Conservative: 2
Best Local Similarity: 54.55% Mismatches: 5
Query Match: 37.76% Indels: 3
DB: 4 Gaps: 1

US-10-014-101B-32 (1-28) x US-09-949-016-13546 (1-163664)

QY 7 LeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSer-----AlaMet 23
Db 121684 TTGCAATACAAAAGGAGGAGGACAAATGGACAAATTGATGTGCACATACGGTGGCTATG 121625

QY 24 IlePro 25
Db 121624 GTCCCT 121619

RESULT 8
US-08-248-466B-4/c
; Sequence 4, Application US/08248466B
; Patent No. 5629182
; GENERAL INFORMATION:
; APPLICANT: CHOPIN, MARIE-CHRISTINE
; APPLICANT: CLUZEL, PIERRE-JEAN
; TITLE OF INVENTION: CLONING OF DNA FRAGMENTS ENCODING A
; TITLE OF INVENTION: MECHANISM FOR RESISTANCE TO BACTERIOPHAGES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/248,466B
; FILING DATE: 24-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/988,959
; FILING DATE: 15-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5629182man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 384-033-0 PCT FWC CIP
```

```
; FILING DATE: 14-SEP-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5629182man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 384-033-0 PCT FWC CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 677 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; US-08-248-466B-4

Alignment Scores:
Pred. No.: 4.4 Length: 677
Score: 55.00 Matches: 12
Percent Similarity: 58.33% Conservative: 2
Best Local Similarity: 50.00% Mismatches: 10
Query Match: 37.41% Indels: 0
DB: 1 Gaps: 0

US-10-014-101B-32 (1-28) x US-08-248-466B-4 (1-677)

QY 5 LeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSerAlaMetIle 24
Db 540 TTAGCCCTTTAAATTCCTTAACTGTCATAAAGTGGAACGTTTTTCTGTAAAGCCTG 481

QY 25 ProGluIleAsp 28
Db 480 CCATTATTGAC 469

RESULT 9
US-08-248-466B-2/c
; Sequence 2, Application US/08248466B
; Patent No. 5629182
; GENERAL INFORMATION:
; APPLICANT: CHOPIN, MARIE-CHRISTINE
; APPLICANT: CLUZEL, PIERRE-JEAN
; TITLE OF INVENTION: CLONING OF DNA FRAGMENTS ENCODING A
; TITLE OF INVENTION: MECHANISM FOR RESISTANCE TO BACTERIOPHAGES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/248,466B
; FILING DATE: 24-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/988,959
; FILING DATE: 15-MAR-1993
; APPLICATION NUMBER: FR 90/11381
; FILING DATE: 14-SEP-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5629182man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 384-033-0 PCT FWC CIP
```

OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (207)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (210)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (213)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (216)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (219)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (225)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (228)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (231)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (234)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (237)  
OTHER INFORMATION: a,g,c or t

Alignment Scores:  
Pred. No.: 4.15 Length: 1602  
Score: 58.00 Matches: 10  
Percent Similarity: 56.00% Conservative: 4  
Best Local Similarity: 40.00% Mismatches: 11  
Query Match: 39.46% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-32 (1-28) x US-09-663-326-10 (1-1602)

QY 4 GlyLeuAlaLeuLeuTyProThrAsnArgAsnLySTrpAspAsnArgMetSerAlaMet 23  
Db 1276 GGCNCNTNAHGNTAYCCNYTAAVARSNATGTGGGAYGAYGNATGNSNGCNCN 1335  
QY 24 IleProGluIleAsp 28  
Db 1336 ACNCNWSNGARGAY 1350

RESULT 5  
US-09-949-016-62602  
Sequence 62602, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 62602  
TYPE: DNA  
ORGANISM: Human

US-09-949-016-62602

Alignment Scores:  
Pred. No.: 3.09 Length: 601  
Score: 55.50 Matches: 12  
Percent Similarity: 63.64% Conservative: 2  
Best Local Similarity: 54.55% Mismatches: 5  
Query Match: 37.76% Indels: 3  
DB: 4 Gaps: 1

US-10-014-101B-32 (1-28) x US-09-949-016-62602 (1-601)

QY 7 LeuLeuTyProThrAsnArgAsnLySTrpAspAsnArgMetSer-----AlaMet 23  
Db 299 TTTCATATACAAAAGGAGGAGGACAAATTTGATGTCACATACGGCTGCTATG 358  
QY 24 IlePro 25  
Db 359 GTCCCT 364

RESULT 6

US-09-949-016-12545/c  
Sequence 12545, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 12545  
LENGTH: 163662  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)-(163662)  
OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-12545

Alignment Scores:

Pred. No.: 4.68e+03 Length: 163662  
Score: 55.50 Matches: 12  
Percent Similarity: 63.64% Conservative: 2  
Best Local Similarity: 54.55% Mismatches: 5  
Query Match: 37.76% Indels: 3  
DB: 4 Gaps: 1

US-10-014-101B-32 (1-28) x US-09-949-016-12545 (1-163662)

QY 7 LeuLeuTyProThrAsnArgAsnLySTrpAspAsnArgMetSer-----AlaMet 23  
Db 121684 TTGCAATACAAAAGGAGGAGGACAAATTTGATGTCACATACGGCTGCTATG 121625  
QY 24 IlePro 25  
Db 121624 GTCCCT 121619

RESULT 7

US-09-949-016-13546/c  
Sequence 13546, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.

NAME/KEY: variation  
LOCATION: (12)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (21)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (24)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (27)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (30)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (33)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (36)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (42)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (48)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (54)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (57)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (60)  
OTHER INFORMATION: a,c,g or t  
NAME/KEY: variation  
LOCATION: (63)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (66)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (69)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (72)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (75)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (78)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (81)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (90)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (93)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (96)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (99)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (105)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation

LOCATION: (108)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (111)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (114)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (117)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (120)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (123)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (126)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (129)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (135)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (141)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (144)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (147)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (153)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (159)  
OTHER INFORMATION: a,c,g or t  
NAME/KEY: variation  
LOCATION: (162)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (165)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (168)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (171)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (174)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (177)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (186)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (195)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (198)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (201)  
OTHER INFORMATION: a,g,c or t  
NAME/KEY: variation  
LOCATION: (204)

/	/	LOCATION: (81)	/
/	/	OTHER INFORMATION: a,g,c or t	/
/	/	FEATURE:	/
/	/	NAME/KEY: variation	/
/	/	LOCATION: (90)	/
/	/	OTHER INFORMATION: a,g,c or t	/
/	/	FEATURE:	/
/	/	NAME/KEY: variation	/
/	/	LOCATION: (93)	/
/	/	OTHER INFORMATION: a,g,c or t	/
/	/	FEATURE:	/
/	/	NAME/KEY: variation	/
/	/	LOCATION: (99)	/
/	/	OTHER INFORMATION: a,g,c or t	/
/	/	FEATURE:	/
/	/	NAME/KEY: variation	/
/	/	LOCATION: (105)	/
/	/	OTHER INFORMATION: a,g,c or t	/
/	/	FEATURE:	/
/	/	NAME/KEY: variation	/
/	/	LOCATION: (108)	/
/	/	OTHER INFORMATION: a,g,c or t	/
/	/	FEATURE:	/
/	/	NAME/KEY: variation	/
/	/	LOCATION: (111)	/
/	/	OTHER INFORMATION: a,g,c or t	/
/	/	FEATURE:	/
/	/	NAME/KEY: variation	/
/	/	LOCATION: (114)	/
/	/	OTHER INFORMATION: a,g,c or t	/
/	/	FEATURE:	/
/	/	NAME/KEY: variation	/
/	/	LOCATION: (117)	/
/	/	OTHER INFORMATION: a,g,c or t	/
/	/	FEATURE:	/
/	/	NAME/KEY: variation	/
/	/	LOCATION: (120)	/
/	/	OTHER INFORMATION: a,g,c or t	/
/	/	FEATURE:	/
/	/	NAME/KEY: variation	/
/	/	LOCATION: (123)	/
/	/	OTHER INFORMATION: a,g,c or t	/
/	/	FEATURE:	/
/	/	NAME/KEY: variation	/
/	/	LOCATION: (126)	/
/	/	OTHER INFORMATION: a,g,c or t	/
/	/	FEATURE:	/
/	/	NAME/KEY: variation	/
/	/	LOCATION: (129)	/
/	/	OTHER INFORMATION: a,g,c or t	/
/	/	FEATURE:	/
/	/	NAME/KEY: variation	/
/	/	LOCATION: (135)	/
/	/	OTHER INFORMATION: a,g,c or t	/
/	/	FEATURE:	/
/	/	NAME/KEY: variation	/
/	/	LOCATION: (141)	/
/	/	OTHER INFORMATION: a,g,c or t	/
/	/	FEATURE:	/
/	/	NAME/KEY: variation	/
/	/	LOCATION: (144)	/
/	/	OTHER INFORMATION: a,g,c or t	/
/	/	FEATURE:	/
/	/	NAME/KEY: variation	/
/	/	LOCATION: (147)	/
/	/	OTHER INFORMATION: a,g,c or t	/
/	/	FEATURE:	/
/	/	NAME/KEY: variation	/
/	/	LOCATION: (153)	/

QY 24 IleProGluLeuAsp 28  
Db 1336 ACGCCGCTCTGAGGAC 1350  
RESULT 2  
US-09-663-326-3  
; Sequence 3, Application US/09663326  
; Patent No. 6617497  
; GENERAL INFORMATION:  
; APPLICANT: Morris Ph.D., Roy O.  
; TITLE OF INVENTION: A CYTOKININ OXIDASE  
; FILE REFERENCE: UMO1490  
; CURRENT APPLICATION NUMBER: US/09/663,326  
; CURRENT FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: 60/054,268  
; PRIOR FILING DATE: 1997-07-30  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1605  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURES:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1605)  
US-09-663-326-3

Alignment Scores:  
Pred. No.: 0.578 Length: 1605  
Score: 63.00 Matches: 11  
Percent Similarity: 64.00% Conservative: 5  
Best Local Similarity: 44.00% Mismatches: 9  
Query Match: 42.86% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101b-32 (1-28) x US-09-663-326-3 (1-1605)

QY 4 GlyLeuAlaLeuLeuTyrProThrAsnArgAsnLysTrpAspAsnArgMetSerAlaMet 23  
Db 1276 GGCCCGCTCATCTCTACCCCTCAACAAATCCATGTGGGACGACGCCCATGTCCGGGGCG 1335  
QY 24 IleProGluLeuAsp 28  
Db 1336 ACGCCGCTCTGAGGAC 1350

RESULT 3  
US-09-124-541-10  
; Sequence 10, Application US/09124541A  
; Patent No. 6229066  
; GENERAL INFORMATION:  
; APPLICANT: Morris Ph.D., Roy O.  
; TITLE OF INVENTION: A CYTOKININ OXIDASE  
; FILE REFERENCE: UMO1490  
; CURRENT APPLICATION NUMBER: US/09/124,541A  
; CURRENT FILING DATE: 1998-07-29  
; EARLIER APPLICATION NUMBER: 60/054,268  
; EARLIER FILING DATE: 1997-07-30  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 10  
; LENGTH: 1602  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURES:  
; NAME/KEY: variation  
; LOCATION: (6)  
; OTHER INFORMATION: a,g,c or t  
; FEATURE:  
; NAME/KEY: variation  
; LOCATION: (9)  
; OTHER INFORMATION: a,g,c or t  
; FEATURE:

; NAME/KEY: variation  
; LOCATION: (12)  
; OTHER INFORMATION: a,g,c or t  
; FEATURE:  
; NAME/KEY: variation  
; LOCATION: (21)  
; OTHER INFORMATION: a,g,c or t  
; FEATURE:  
; NAME/KEY: variation  
; LOCATION: (24)  
; OTHER INFORMATION: a,g,c or t  
; FEATURE:  
; NAME/KEY: variation  
; LOCATION: (27)  
; OTHER INFORMATION: a,g,c or t  
; FEATURE:  
; NAME/KEY: variation  
; LOCATION: (30)  
; OTHER INFORMATION: a,g,c or t  
; FEATURE:  
; NAME/KEY: variation  
; LOCATION: (33)  
; OTHER INFORMATION: a,g,c or t  
; FEATURE:  
; NAME/KEY: variation  
; LOCATION: (36)  
; OTHER INFORMATION: a,g,c or t  
; FEATURE:  
; NAME/KEY: variation  
; LOCATION: (42)  
; OTHER INFORMATION: a,g,c or t  
; FEATURE:  
; NAME/KEY: variation  
; LOCATION: (48)  
; OTHER INFORMATION: a,g,c or t  
; FEATURE:  
; NAME/KEY: variation  
; LOCATION: (54)  
; OTHER INFORMATION: a,g,c or t  
; FEATURE:  
; NAME/KEY: variation  
; LOCATION: (57)  
; OTHER INFORMATION: a,g,c or t  
; FEATURE:  
; NAME/KEY: variation  
; LOCATION: (60)  
; OTHER INFORMATION: a,g or t  
; FEATURE:  
; NAME/KEY: variation  
; LOCATION: (63)  
; OTHER INFORMATION: a,g,c or t  
; FEATURE:  
; NAME/KEY: variation  
; LOCATION: (66)  
; OTHER INFORMATION: a,g,c or t  
; FEATURE:  
; NAME/KEY: variation  
; LOCATION: (69)  
; OTHER INFORMATION: a,g,c or t  
; FEATURE:  
; NAME/KEY: variation  
; LOCATION: (72)  
; OTHER INFORMATION: a,g,c or t  
; FEATURE:  
; NAME/KEY: variation  
; LOCATION: (75)  
; OTHER INFORMATION: a,g,c or t  
; FEATURE:  
; NAME/KEY: variation  
; LOCATION: (78)  
; OTHER INFORMATION: a,g,c or t  
; FEATURE:  
; NAME/KEY: variation

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 18, 2005, 09:41:47 ; Search time 469 seconds  
(without alignments)  
97.688 Million cell updates/sec

Title: US-10-014-101B-32  
Perfect score: 147  
Sequence: 1 SASGALLYPTNRKNWDRMSAMPEID 28

Scoring table: BLOSUM62 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10014101/runat\_16022005\_075935\_8326/app\_query.fasta\_1.199  
-DB=Issued Patents NA -OPMT=fastcap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40  
-MODE=LOCAL -OUTFWT=pco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10014101.cgn2\_1\_54 @runat\_16022005\_075935\_8326 -NCPU=6 -ICPU=3  
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	63	42.9	1605	US-09-124-541-3	Sequence 3, Appli
2	63	42.9	1605	US-09-663-326-3	Sequence 3, Appli
3	58	39.5	1602	US-09-124-541-10	Sequence 10, Appl
4	58	39.5	1602	US-09-663-326-10	Sequence 10, Appl
5	55.5	37.8	1601	US-09-949-016-62602	Sequence 12545, A
6	55.5	37.8	163662	US-09-949-016-12545	Sequence 12545, A
7	55.5	37.8	163664	US-09-949-016-13546	Sequence 13546, A
8	55	37.4	677	US-08-248-466B-4	Sequence 4, Appli
9	55	37.4	2763	US-08-248-466B-2	Sequence 2, Appli
10	55	37.4	32616	US-09-949-016-13407	Sequence 13407, A
11	54	36.7	1031	US-09-023-655-543	Sequence 543, App
12	54	36.7	28555	US-09-949-016-13046	Sequence 13046, A

C 13	53.5	36.4	3306	1	US-08-261-206A-71	Sequence 71, Appl
C 14	52	35.4	5198	1	US-08-123-761A-1	Sequence 1, Appli
C 15	52	35.4	84870	4	US-09-949-016-17547	Sequence 17547, A
C 16	51	34.7	1893	4	US-09-657-252-3	Sequence 3, Appli
C 17	51	34.7	53394	4	US-09-949-016-15817	Sequence 15817, A
C 18	51	34.7	53394	4	US-09-949-016-15818	Sequence 15818, A
C 19	51	34.7	53394	4	US-09-949-016-15819	Sequence 15819, A
C 20	51	34.7	53394	4	US-09-949-016-15820	Sequence 15820, A
C 21	51	34.7	62909	4	US-09-596-002-32	Sequence 32, Appl
C 22	50	34.0	510	4	US-09-902-540-6994	Sequence 6994, Ap
C 23	50	34.0	547	4	US-09-663-600A-137	Sequence 137, App
C 24	50	34.0	648	4	US-09-663-600A-43	Sequence 43, Appl
C 25	50	34.0	1008	4	US-09-248-798A-738	Sequence 738, App
C 26	50	34.0	2085	4	US-09-149-718-1	Sequence 1, Appli
C 27	50	34.0	2388	4	US-09-902-540-611	Sequence 611, App
C 28	50	34.0	390416	4	US-09-949-016-16923	Sequence 16923, A
C 29	49.5	33.7	6439	4	US-09-902-540-813	Sequence 813, App
C 30	49.5	33.7	27579	4	US-09-949-016-13465	Sequence 13465, A
C 31	49.5	33.7	14115	4	US-09-949-016-17490	Sequence 17490, A
C 32	49	33.3	120	1	US-08-458-518-21	Sequence 21, Appl
C 33	49	33.3	350	4	US-09-270-767-1938	Sequence 1938, Ap
C 34	49	33.3	350	4	US-09-270-767-17220	Sequence 17220, A
C 35	49	33.3	858	4	US-09-543-681A-925	Sequence 925, App
C 36	49	33.3	863	3	US-08-728-603-18	Sequence 18, Appl
C 37	49	33.3	948	4	US-09-489-039A-2926	Sequence 2926, Ap
C 38	49	33.3	1413	4	US-09-543-681A-3221	Sequence 3221, Ap
C 39	49	33.3	2601	3	US-09-221-017B-1014	Sequence 1014, Ap
C 40	49	33.3	32207	2	US-08-770-379-20	Sequence 20, Appl
C 41	49	33.3	32207	3	US-08-757-669A-20	Sequence 20, Appl
C 42	49	33.3	32207	3	US-09-230-371A-20	Sequence 20, Appl
C 43	49	33.3	192506	4	US-09-949-016-15830	Sequence 15830, A
C 44	48.5	33.0	137949	4	US-09-949-016-12196	Sequence 12196, A
C 45	48.5	33.0	137956	4	US-09-949-016-17260	Sequence 17260, A

ALIGNMENTS

RESULT 1  
US-09-124-541-3  
; Sequence 3, Application US/09124541A  
; Patent No. 6229086  
; GENERAL INFORMATION:  
; APPLICANT: Morris Ph.D., Roy O.  
; TITLE OF INVENTION: A CYTOKININ OXIDASE  
; FILE REFERENCE: UMO1490  
; CURRENT APPLICATION NUMBER: US/09/124,541A  
; CURRENT FILING DATE: 1998-07-29  
; EARLIER APPLICATION NUMBER: 60/054,268  
; EARLIER FILING DATE: 1997-07-30  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1605  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1605)  
US-09-124-541-3

Alignment Scores:  
Pred. No.: 0.578 Length: 1605  
Score: 63.00 Matches: 11  
Percent Similarity: 64.00% Conservative: 5  
Best Local Similarity: 44.00% Mismatches: 9  
Query Match: 42.86% Indels: 0  
DB: 3 Gaps: 0

US-10-014-101B-32 (1-28) x US-09-124-541-3 (1-1605)

Qy 4 GlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSerAlaMet 23  
Db 1276 GGCCGCTCATGCTCTACCCCTCACAAATCCATGTGGGACGACGCGATGTCGGCGGCG 1335

```

XX (PHAA ) PHARMACIA CORP.
XX
XX Huang S, Crossland LD, Cheikh N, Morris RO;
XX
XX WPI; 2003-897983/82.
XX GENBANK; AY091158.
XX
XX Producing plants characterized by reversible male-sterility, useful for
XX PT maintaining male sterility in plants, by transforming a plant cell with a
XX PT nucleic acid construct containing a polynucleotide encoding a cytokinin
XX PT oxide.
XX
XX PS Disclosure; SEQ ID NO 21; 33pp; English.
XX
XX CC The invention relates to a method for producing a plant characterised by
XX CC reversible male-sterility which involves transforming a plant cell with a
XX CC nucleic acid construct containing a polynucleotide encoding a cytokinin
XX CC oxidase. The method is useful for producing reversible male-sterility in
XX CC transgenic plants, or for maintaining male sterility in plants. The
XX CC method reduces the expense of seed production for existing hybrid plants
XX CC such as corn, but also makes it possible to produce hybrid varieties of
XX CC traditionally non-hybrid crops. The method is also useful for introducing
XX CC economically valuable traits from plants having undesirable production
XX CC characteristics into plants having desirable characteristics. The present
XX CC sequence is Arabidopsis thaliana cytokinin oxidase (CKX1) DNA. This
XX CC sequence is used to illustrate the method of the invention.
XX
XX SQ Sequence 1873 BP; 495 A; 385 C; 482 G; 511 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0.189 Length: 1873
Score: 69.00 Matches: 12
Percent Similarity: 75.00% Conservative: 6
Best Local Similarity: 50.00% Mismatches: 6
Query Match: 46.94% Indels: 0
DB: 10 Gaps: 0

US-10-014-101B-32 (1-28) x ADH61280 (1-1873)

Qy 3 Ser(GlyLeuAlaLeuLeuTyrProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22
Db 1340 AATGGTCCATGCTTGTGTACCCCACTCTTGGCAGACGAGTGGGAGTACGACGCTCGTG 1399
Qy 23 MetIleProGlu 26
Db 1400 GTTATACCGGA 1411

```

Search completed: February 18, 2005, 10:16:45  
Job time : 264 secs

XX PD 28-AUG-2003.  
 XX XX  
 XX PF 20-DEC-2002; 2002US-00326184.  
 XX XX  
 XX PR 20-DEC-2001; 2001US-0343129P.  
 XX PA (PHAA ) PHARMACIA CORP.  
 XX PI Huang S, Crossland LD, Cheikh N, Morris RO;  
 XX PI WPI; 2003-897983/82.  
 XX DR GENBANK; AF540382.  
 XX PT Producing plants characterized by reversible male-sterility, useful for  
 PT maintaining male sterility in plants, by transforming a plant cell with a  
 PT nucleic acid construct containing a polynucleotide encoding a cytokinin  
 PT oxidase.  
 XX PS Disclosure; SEQ ID NO 22; 33pp; English.  
 XX CC The invention relates to a method for producing a plant characterised by  
 CC reversible male-sterility which involves transforming a plant cell with a  
 CC nucleic acid construct containing a polynucleotide encoding a cytokinin  
 CC oxidase. The method is useful for producing reversible male-sterility in  
 CC transgenic plants, or for maintaining male sterility in plants. The  
 CC method reduces the expense of seed production for existing hybrid plants  
 CC such as corn, but also makes it possible to produce hybrid varieties of  
 CC economically valuable traits from plants having undesirable production  
 CC characteristics into plants having desirable characteristics. The present  
 CC sequence is barley cytokinin oxidase (CKX1) DNA. This sequence is used to  
 CC illustrate the method of the invention.  
 XX SQ Sequence 1857 BP; 521 A; 454 C; 437 G; 445 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 0.0173 Length: 1857  
 Score: 75.00 Matches: 12  
 Percent Similarity: 75.00% Conservative: 6  
 Best Local Similarity: 75.00% Mismatches: 6  
 Query Match: 51.02% Indels: 0  
 DB: 10 Gaps: 0  
 US-10-014-101B-32 (1-28) x ADH61281 (1-1857)  
 QY 3 SerGlyLeuAlaLeuLeuTyPProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22  
 Db 1298 AATGGTCCCATATTGCTCTACCCAGTGAGAGAGTCCAGATGGGACACCGACGTCAGTG 1357  
 QY 23 MetIleProGlu 26  
 Db 1358 GTCATACCCAGAT 1369  
 RESULT 39  
 ADH61270  
 ID ADH61270 standard; DNA; 1575 BP.  
 XX AC ADH61270;  
 XX DT 25-MAR-2004 (first entry)  
 XX DE Arabidopsis thaliana cytokinin oxidase (CKX1) DNA #4.  
 XX KW Male-sterility; cytokinin oxidase; transgenic plant; seed production;  
 XX KW CKX1; mouse-ear cross; gene; ds.  
 XX OS Arabidopsis thaliana.  
 XX PN US2003163847-A1.  
 XX PD 28-AUG-2003.  
 XX PF 20-DEC-2002; 2002US-00326184.  
 XX PR 20-DEC-2001; 2001US-0343129P.

PF 20-DEC-2002; 2002US-00326184.  
 XX XX  
 XX PR 20-DEC-2001; 2001US-0343129P.  
 XX XX  
 XX PA (PHAA ) PHARMACIA CORP.  
 XX PI Huang S, Crossland LD, Cheikh N, Morris RO;  
 XX PI WPI; 2003-897983/82.  
 XX DR GENBANK; AF303981.  
 XX PT Producing plants characterized by reversible male-sterility, useful for  
 PT maintaining male sterility in plants, by transforming a plant cell with a  
 PT nucleic acid construct containing a polynucleotide encoding a cytokinin  
 PT oxidase.  
 XX PS Claim 44; SEQ ID NO 11; 33pp; English.  
 XX CC The invention relates to a method for producing a plant characterised by  
 CC reversible male-sterility which involves transforming a plant cell with a  
 CC nucleic acid construct containing a polynucleotide encoding a cytokinin  
 CC oxidase. The method is useful for producing reversible male-sterility in  
 CC transgenic plants, or for maintaining male sterility in plants. The  
 CC method reduces the expense of seed production for existing hybrid plants  
 CC such as corn, but also makes it possible to produce hybrid varieties of  
 CC economically valuable traits from plants having undesirable production  
 CC characteristics into plants having desirable characteristics. The present  
 CC sequence is Arabidopsis thaliana cytokinin oxidase (CKX1) DNA. This  
 CC sequence is used to illustrate the method of the invention.  
 XX SQ Sequence 1575 BP; 390 A; 328 C; 451 G; 406 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 0.151 Length: 1575  
 Score: 69.00 Matches: 12  
 Percent Similarity: 75.00% Conservative: 6  
 Best Local Similarity: 50.00% Mismatches: 6  
 Query Match: 46.94% Indels: 0  
 DB: 10 Gaps: 0  
 US-10-014-101B-32 (1-28) x ADH61270 (1-1575)  
 QY 3 SerGlyLeuAlaLeuLeuTyPProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22  
 Db 1225 AATGGTCCCATCTGTGTACCCACTCTTGGAGACAGTGGGATGATCGACGTCCTG 1284  
 QY 23 MetIleProGlu 26  
 Db 1285 GTTATACCGGAA 1296  
 RESULT 40  
 ADH61280  
 ID ADH61280 standard; DNA; 1873 BP.  
 XX AC ADH61280;  
 XX DT 25-MAR-2004 (first entry)  
 XX DE Arabidopsis thaliana cytokinin oxidase (CKX1) DNA #7.  
 XX KW Male-sterility; cytokinin oxidase; transgenic plant; seed production;  
 XX KW CKX1; mouse-ear cross; gene; ds.  
 XX OS Arabidopsis thaliana.  
 XX PN US2003163847-A1.  
 XX PD 28-AUG-2003.  
 XX PF 20-DEC-2002; 2002US-00326184.  
 XX PR 20-DEC-2001; 2001US-0343129P.

KW embryo size; cotyledon size; transgenic plant; herbicide; gene; ss.  
 XX Arabidopsis thaliana.  
 OS W02003050287-A2.  
 XX 19-JUN-2003.  
 XX 10-DEC-2002; 2002MO-EP013990.  
 XX 10-DEC-2001; 2001US-00014101.  
 XX (SCHM/) SCHMULLING T.  
 PA (WERN/) WERNER T.  
 XX Schmulling T, Werner T;  
 XX WPI; 2003-541577/51.  
 XX Stimulating root growth, enhancing lateral or adventitious root formation  
 PT or altering root geotropism comprises increasing plant cytokinin oxidase  
 PT levels or other protein or nucleic acid that reduces active cytokinins in  
 PT a plant.  
 XX Claim 3; Page 174-175; 177pp; English.  
 XX The present invention relates to a method for stimulating root growth or  
 CC enhancing the formation of lateral or adventitious roots or altering root  
 CC geotropism, which comprises increasing in a plant or plant part the level  
 CC of a plant cytokinin oxidase or other protein that reduces the level of  
 CC active cytokinins in a plant or plant part. Cytokinin oxidase protein and  
 CC coding sequences from Arabidopsis thaliana are also provided. The method  
 CC is useful in modifying plant morphological, biochemical and physiological  
 CC properties, such as in modifying the initiation, stimulation or  
 CC enhancement of root growth, adventitious root formation, lateral root  
 CC formation, root geotropism, shoot growth, apical dominance, branching,  
 CC timing of senescence, timing of flowering, flower formation, seed  
 CC development and/or seed yield. The present sequence is a coding sequence  
 CC shown in the invention  
 XX SQ Sequence 1620 BP; 419 A; 380 C; 411 G; 410 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 0.00653 Length: 1620  
 Score: 77.00 Matches: 13  
 Percent Similarity: 75.00% Conservative: 5  
 Best Local Similarity: 54.17% Mismatches: 6  
 Query Match: 52.38% Indels: 0  
 DB: 9 Gaps: 0  
 US-10-014-101B-32 (1-28) x ACC85301 (1-1620)  
 QY 3 SerGlyLeuAlaLeuLeuTyrProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22  
 Db 1222 AGTGGCCCTATTCTTATCTACCCCATGACAAAGACAAATGGGACGAGGAGCTCAGCC 1281  
 QY 23 MetIleProGlu 26  
 Db 1282 GTGACGCCGGAT 1293  
 RESULT 37  
 ADH61271  
 ID ADH61271 standard; DNA; 1623 BP.  
 XX ADH61271;  
 AC ADH61271;  
 XX 25-MAR-2004 (first entry)  
 XX Arabidopsis thaliana cytokinin oxidase (CKX1) DNA #5.  
 DE Male-sterility; cytokinin oxidase; transgenic plant; seed production;  
 XX CKX1; mouse-ear cress; gene; ds.  
 KW Male-sterility; cytokinin oxidase; transgenic plant; seed production;  
 XX CKX1; mouse-ear cress; gene; ds.

OS Arabidopsis thaliana.  
 XX US2003163847-A1.  
 XX 28-AUG-2003.  
 XX 20-DEC-2002; 2002US-00326184.  
 XX 20-DEC-2001; 2001US-0343129P.  
 XX (PHAA ) PHARMACIA CORP.  
 XX Huang S, Crossland LD, Cheikh N, Morris RO;  
 XX WPI; 2003-897983/82.  
 XX GENBANK; AF303982.  
 XX Producing plants characterized by reversible male-sterility, useful for  
 PT maintaining male sterility in plants, by transforming a plant cell with a  
 PT nucleic acid construct containing a polynucleotide encoding a cytokinin  
 PT oxidase.  
 XX Claim 45; SEQ ID NO 12; 33pp; English.  
 XX The invention relates to a method for producing a plant characterised by  
 CC reversible male-sterility which involves transforming a plant cell with a  
 CC nucleic acid construct containing a polynucleotide encoding a cytokinin  
 CC oxidase. The method is useful for producing reversible male-sterility in  
 CC transgenic plants, or for maintaining male sterility in plants. The  
 CC method reduces the expense of seed production for existing hybrid plants  
 CC such as corn, but also makes it possible to produce hybrid varieties of  
 CC traditionally non-hybrid crops. The method is also useful for introducing  
 CC economically valuable traits from plants having undesirable production  
 CC characteristics into plants having desirable characteristics. The present  
 CC sequence is Arabidopsis thaliana cytokinin oxidase (CKX1) DNA. This  
 CC sequence is used to illustrate the method of the invention.  
 XX SQ Sequence 1623 BP; 421 A; 382 C; 412 G; 408 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 0.00655 Length: 1623  
 Score: 77.00 Matches: 13  
 Percent Similarity: 75.00% Conservative: 5  
 Best Local Similarity: 54.17% Mismatches: 6  
 Query Match: 52.38% Indels: 0  
 DB: 10 Gaps: 0  
 US-10-014-101B-32 (1-28) x ADH61271 (1-1623)  
 QY 3 SerGlyLeuAlaLeuLeuTyrProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22  
 Db 1225 AGTGGCCCTATTCTTATCTACCCCATGACAAAGACAAATGGGACGAGGAGCTCAGCC 1284  
 QY 23 MetIleProGlu 26  
 Db 1285 GTGACGCCGGAT 1296  
 RESULT 38  
 ADH61281  
 ID ADH61281 standard; DNA; 1857 BP.  
 XX ADH61281;  
 AC ADH61281;  
 XX 25-MAR-2004 (first entry)  
 XX Barley cytokinin oxidase (CKX1) DNA #1.  
 DE Male-sterility; cytokinin oxidase; transgenic plant; seed production;  
 XX CKX1; barley; ds; gene.  
 KW Male-sterility; cytokinin oxidase; transgenic plant; seed production;  
 XX CKX1; barley; ds; gene.  
 OS Hordeum vulgare.  
 XX US2003163847-A1.

XX 19-JUN-2003.  
PD 10-DEC-2002; 2002WO-EP013990.  
XX 10-DEC-2001; 2001US-00014101.  
XX (SCHM/) SCHMULLING T.  
PA (WERN/) WERNER T.  
XX Schmullling T, Werner T;  
XX WPI; 2003-541577/51.  
XX Stimulating root growth, enhancing lateral or adventitious root formation  
PT or altering root geotropism comprises increasing plant cytokinin oxidase  
PT levels or other protein or nucleic acid that reduces active cytokinins in  
PT a plant.  
XX Claim 3; Page 170-171; 177pp; English.  
XX The present invention relates to a method for stimulating root growth or  
CC enhancing the formation of lateral or adventitious roots or altering root  
CC geotropism, which comprises increasing in a plant or plant part the level  
CC of a plant cytokinin oxidase or other protein that reduces the level of  
CC active cytokinins in a plant or plant part. Cytokinin oxidase protein and  
CC coding sequences from Arabidopsis thaliana are also provided. The method  
CC is useful in modifying plant morphological, biochemical and physiological  
CC properties, such as in modifying the initiation, stimulation or  
CC enhancement of root growth, adventitious root formation, lateral root  
CC formation, root geotropism, shoot growth, apical dominance, branching,  
CC timing of senescence, timing of flowering, flower formation, seed  
CC development and/or seed yield. The present sequence is a coding sequence  
CC shown in the invention  
XX Sequence 1611 BP; 416 A; 379 C; 409 G; 407 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 0.00648 Length: 1611  
Score: 77.00 Matches: 13  
Percent Similarity: 75.00% Conservatives: 5  
Best Local Similarity: 54.17% Mismatches: 6  
Query Match: 52.38% Indels: 0  
DB: Gaps: 0  
US-10-014-101b-32 (1-28) x ACC85297 (1-1611)  
QY 3 SerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22  
Db 1213 AGTGGCCCTATTCTTATCTACCCCATGACAAAGACAAATGGGACGAGGAGCTCAGCC 1272  
QY 23 MetIleProGlu 26  
Db 1273 GTGACGCCGGAT 1284  
RESULT 35  
ABK28632  
ID ABK28632 standard; cDNA; 1620 BP.  
XX AC ABK28632;  
XX 09-APR-2002 (first entry)  
XX cDNA encoding A. thaliana cytokinin oxidase AtCKX5 (long version).  
XX Arabidopsis thaliana; plant; cytokinin oxidase; transgenic plant;  
XX root growth; lateral root; adventitious root; root geotropism; herbicide;  
XX root meristem; shoot meristem; leaf senescence; parthenocarpy; gene; ss.  
XX Arabidopsis thaliana.  
XX OS  
XX WO200196580-A2.  
XX PN  
XX

PD 20-DEC-2001.  
XX 18-JUN-2001; 2001WO-EP006833.  
XX 16-JUN-2000; 2000EP-00870132.  
PR 27-DEC-2000; 2000US-0258413P.  
PR 16-MAR-2001; 2001EP-00870053.  
XX (SCHM/) SCHMULLING T.  
PA (WERN/) WERNER T.  
XX Schmullling T, Werner T;  
XX WPI; 2002-130736/17.  
DR P-PSDB; AAU81974.  
XX Polynucleotide encoding novel plant protein having cytokinin oxidase  
PT activity and the protein useful for stimulating root growth, enhancing  
PT the formation of lateral or adventitious roots, altering root geotropism.  
XX Claim 3; Page 151; 154pp; English.  
XX The invention relates to an isolated polynucleotide (I) encoding a novel  
CC plant protein (II) having cytokinin oxidase activity. (I) is useful for  
CC production of transgenic plants, plant cells or tissues; for production  
CC of altered plants, plant cell or tissues; and for effecting the  
CC expression of (II) where (I) is operably linked to one or more control  
CC sequences. The methods further comprises regenerating a plant from the  
CC plant cell. (I) and (II) are useful for stimulating root growth;  
CC enhancing the formation of lateral or adventitious roots; altering root  
CC geotropism, leading to an increase in yield; and for screening growth  
CC promoting chemical of herbicides. (I) is useful for increasing the size  
CC of the root meristem; increasing root size; increasing the size of the  
CC shoot meristem; delaying leaf senescence and altering leaf senescence;  
CC increasing leaf thickness; reducing or increasing the vessel size;  
CC inducing parthenocarpy; improving standability of the seedlings;  
CC increasing branching and for improving lodging resistance. Antibody (III)  
CC to (II) is useful for identifying and obtaining proteins interacting with  
CC (II) comprising a screening assay, preferably a two-hybrid screening  
CC assay. ABK28606-ABK28633 represent Arabidopsis thaliana cytokinin oxidase  
CC coding sequences and PCR primers of the invention  
XX Sequence 1620 BP; 419 A; 380 C; 411 G; 410 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 0.00653 Length: 1620  
Score: 77.00 Matches: 13  
Percent Similarity: 75.00% Conservatives: 5  
Best Local Similarity: 54.17% Mismatches: 6  
Query Match: 52.38% Indels: 0  
DB: Gaps: 0  
US-10-014-101b-32 (1-28) x ABK28632 (1-1620)  
QY 3 SerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22  
Db 1222 AGTGGCCCTATTCTTATCTACCCCATGACAAAGACAAATGGGACGAGGAGCTCAGCC 1281  
QY 23 MetIleProGlu 26  
Db 1282 GTGACGCCGGAT 1293  
RESULT 36  
ACC85301  
ID ACC85301 standard; cDNA; 1620 BP.  
XX AC ACC85301;  
XX 18-SEP-2003 (first entry)  
XX Arabidopsis cytokinin oxidase-like protein 5 cDNA #2.  
XX Maize; root growth; root geotropism; cytokinin oxidase; seed size;  
XX



Query Match: 53.74% Indels: 0  
DB: 13 Gaps: 0

US-10-014-101B-32 (1-28) x ACN54881 (1-482)

QY 3 SerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22  
107 AGTGGCCCAATTCATCTATCCATGACACAAATGGACCATAGGAGCTCCGTG 166

QY 23 MetIleProGlu 26  
DB 167 GTGACACCGGAT 178

RESULT 31  
ACN55634/c  
ID ACN55634 standard; cDNA; 484 BP.  
XX AC ACN55634;  
XX 02-DEC-2004 (first entry)  
XX Cotton androecium tissue EST Clone ID: LIB3828-026-Q6-N6-E7, SEQ:10415.  
XX Cotton; plant; EST; expressed sequence tag; transgenic plant; androecium;  
KW variety Nucotton33B; library LIB3828; molecular tag; molecular marker;  
KW genetic mapping; molecular mapping; seed germination; plant growth;  
KW plant quality; plant yield; plant breeding; tissue printing; ss.  
XX Gossypium hirsutum.  
XX US2004123340-A1.  
XX 24-JUN-2004.  
XX 12-DEC-2001; 2001US-00021323.  
XX 14-DEC-2000; 2000US-0255619P.  
XX (DEIK/) DEIKMAN J.  
PA (FENG/) FENG P C C.  
PA (FING/) FINCHER K L.  
PA (ZIEG/) ZIEGLER T E.  
XX Deikman J, Feng PCC, Fincher KL, Ziegler TE;  
XX WPI; 2004-479808/45.  
XX New isolated nucleic acid molecule that encodes a plant protein or its  
PT fragment, useful for isolating a variety of agronomically significant  
PT genes associated with plant growth, quality or yield, and as molecular  
PT tags to map genes.  
XX Claim 1; SEQ ID NO 10415; 34pp; English.

The invention relates to 17880 cotton expressed sequence tags (ESTs;  
CC ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated  
CC from primed or non-primed seeds from variety DP50B, mature seeds from  
CC variety Coker 312 Boswell 96 Field, and androecium tissue, gynoeium  
CC tissue, developing fibres, carpel walls and septa from variety  
CC Nucotton33B. The invention also relates to substantially purified  
CC proteins or their fragments encoded by nucleic acid molecules of the  
CC invention, and to transformed plants having a nucleic acid construct  
CC comprising a nucleic acid of the invention. The cotton ESTs are useful as  
CC molecular tags to isolate genetic regions, to isolate genes, to map  
CC genes, to determine gene function and to determining whether genes are  
CC members of a particular gene family. The nucleic acid molecules may be  
CC used for isolating a variety of agronomically significant genes  
CC associated with plant growth, quality, yield, and could also serve as  
CC links in metabolic and catabolic pathways. The nucleic acid molecules are  
CC also useful for identifying genes important in initiating and maintaining  
CC seed germination or that may be used to mitigate stresses encountered  
CC during seed germination. The ESTs additionally enable the acquisition of  
CC promoters and cis-regulatory elements which will be useful to express

CC agronomically significant genes in these tissues and/or other tissues,  
CC and also permits the acquisition of molecular markers useful in breeding  
CC schemes, genetic and molecular mapping, and in cloning of agronomically  
CC significant genes. The nucleic acid molecules are further useful for  
CC detecting the expression level or pattern of a protein or mRNA and for  
CC detecting the presence or quantity of a protein by tissue printing. The  
CC present sequence represents a specifically claimed EST isolated from a  
CC cotton variety Nucotton33B androecium tissue cDNA library (LIB3828). The  
CC sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from the US  
CC patent office at seqdata.uspto.gov/sequence.html?DocID=US20040123340  
XX SQ Sequence 484 BP; 121 A; 121 C; 102 G; 139 T; 0 U; 1 Other;

Alignment Scores:  
Pred. No.: 0.00618 Length: 484  
Score: 79.00 Matches: 13  
Percent Similarity: 75.00% Conservative: 5  
Best Local Similarity: 54.17% Mismatches: 6  
Query Match: 53.74% Indels: 0  
DB: 13 Gaps: 0

US-10-014-101B-32 (1-28) x ACN55634 (1-484)

QY 3 SerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22  
DB 376 AGTGGCCCAATTCATCTATCCATGACACAAATGGACCATAGGAGCTCCGTG 317

QY 23 MetIleProGlu 26  
DB 316 GTGACACCGGAT 305

RESULT 32  
ADR61296  
ID ADR61296 standard; cDNA; 788 BP.  
XX ADR61296;  
XX 02-DEC-2004 (first entry)  
XX Cotton cDNA sequence, SEQ ID 2077.  
XX Cotton; ss; plant; cold tolerance; growth rate; cell cycle pathway;  
KW drought tolerance; plant disease resistance; galactomannan; lignin;  
KW plant growth regulator; heat tolerance; herbicide tolerance;  
KW homologous recombination; extreme osmotic condition tolerance;  
KW pathogen resistance; pest resistance; yield; photosynthesis; seed oil;  
KW stress resistance.  
XX Gossypium hirsutum.  
XX US2004181830-A1.  
XX 16-SEP-2004.  
XX 29-JAN-2004; 2004US-00767795.  
XX 07-MAY-2001; 2001US-00849529.  
XX 12-DEC-2001; 2001US-00021323.  
XX (KOVA/) KOVALIC D K.  
PA (ZHOU/) ZHOU Y.  
PA (CAOY/) CAO Y.  
XX Kovalic DK, Zhou Y, Cao Y;  
XX WPI; 2004-667718/65.  
XX New recombinant nucleic acid molecules and polypeptides from Gossypium  
PT hirsutum, useful for producing plants with improved biological  
PT characteristics (e.g. improved plant cold or drought tolerance).  
XX Claim 1; SEQ ID NO 2077; 14pp; English.

AC ACC85298;  
DT 18-SEP-2003 (first entry)  
XX  
DE Arabidopsis cytokinin oxidase-like protein 6 cDNA.  
XX  
KW Maize; root growth; root geotropism; cytokinin oxidase; seed size;  
KW embryo size; cotyledon size; transgenic plant; herbicide; gene; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
PN WO2003050287-A2.  
XX  
PD 19-JUN-2003.  
XX  
PF 10-DEC-2002; 2002WO-EP013990.  
XX  
PR 10-DEC-2001; 2001US-00014101.  
XX  
PA (SCHM/) SCHMULLING T.  
PA (WERN/) WERNER T.  
XX  
PI Schmullling T, Werner T;  
XX  
PI WPI; 2003-541577/51.  
XX  
PT Stimulating root growth, enhancing lateral or adventitious root formation  
PT or altering root geotropism comprises increasing plant cytokinin oxidase  
PT levels or other protein or nucleic acid that reduces active cytokinins in  
PT a plant.  
XX  
PS Claim 2; Page 171-172; 177pp; English.  
XX  
CC The present invention relates to a method for stimulating root growth or  
CC enhancing the formation of lateral or adventitious roots or altering root  
CC geotropism, which comprises increasing in a plant or plant part the level  
CC of a plant cytokinin oxidase or other protein that reduces the level of  
CC active cytokinins in a plant or plant part. Cytokinin oxidase protein and  
CC coding sequences from Arabidopsis thaliana are also provided. The method  
CC is useful in modifying plant morphological, biochemical and physiological  
CC properties, such as in modifying the initiation, stimulation or  
CC enhancement of root growth, adventitious root formation, lateral root  
CC formation, root geotropism, shoot growth, apical dominance, branching,  
CC timing of senescence, timing of flowering, flower formation, seed  
CC development and/or seed yield. The present sequence is a coding sequence  
CC shown in the invention  
XX  
SQ Sequence 1515 BP; 456 A; 327 C; 378 G; 354 T; 0 U; 0 Other;  
  
Alignment Scores:  
Pred. No.: 0.00182 Length: 1515  
Score: 80.00 Matches: 12  
Percent Similarity: 80.77% Conservative: 9  
Best Local Similarity: 46.15% Mismatches: 5  
Query Match: 54.42% Indels: 0  
DB: 9 Gaps: 0  
  
US-10-014-101B-32 (1-28) x ACC85298 (1-1515)  
  
QY 1 SerAlaSerGlyLeuAlaLeuLeuTyrProThrAsnArgAsnLysTrpAspAsnArgMet 20  
Db 1150 ACNAGCAACGGCCAGTCATCGTCTACCCGTAACAAATCAAGTGGGCAATCAACA 1209  
QY 21 SerAlaMetIleProGlu 26  
Db 1210 TCAGCAGTAACACCGGAG 1227  
  
RESULT 30  
ACN54881  
ID ACN54881 standard; cDNA; 482 BP.  
XX  
AC ACN54881;  
XX

DT 02-DEC-2004 (first entry)  
XX  
DE Cotton androecium tissue EST Clone ID: LIB3828-026-Q6-K6-E7, SEQ:9662.  
XX  
KW Cotton; plant; EST; expressed sequence tag; transgenic plant; androecium;  
KW variety Nucotton33B; library LIB3828; molecular tag; molecular marker;  
KW genetic mapping; molecular mapping; seed germination; plant growth;  
KW plant quality; plant yield; plant breeding; tissue printing; ss.  
XX  
OS Gossypium hirsutum.  
XX  
PN US2004123340-A1.  
XX  
PD 24-JUN-2004.  
XX  
PF 12-DEC-2001; 2001US-00021323.  
XX  
PR 14-DEC-2000; 2000US-0255619P.  
XX  
PA (DEIK/) DEIKMAN J.  
PA (FENG/) FENG P C C.  
PA (FINC/) FINCHER K L.  
PA (ZIEG/) ZIEGLER T E.  
XX  
PI Deikman J, Feng PCC, Fincher KL, Ziegler TE;  
XX  
PI WPI; 2004-479808/45.  
XX  
DR New isolated nucleic acid molecule that encodes a plant protein or its  
XX fragment, useful for isolating a variety of agronomically significant  
XX genes associated with plant growth, quality or yield, and as molecular  
XX tags to map genes.  
XX  
PS Claim 1; SEQ ID NO 9662; 34pp; English.  
XX  
CC The invention relates to 17880 cotton expressed sequence tags (ESTs;  
CC ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated  
CC from primed or non-primed seeds from variety DF50B, mature seeds from  
CC variety Coker 312 Boswell 96 Field, and androecium tissue, gynoeceum  
CC tissue, developing fibres, carpel walls and septa from variety  
CC Nucotton33B. The invention also relates to substantially purified  
CC proteins or their fragments encoded by nucleic acid molecules of the  
CC invention, and to transformed plants having a nucleic acid construct  
CC comprising a nucleic acid of the invention. The cotton ESTs are useful as  
CC molecular tags to isolate genetic regions, to isolate genes, to map  
CC genes, to determine gene function and to determining whether genes are  
CC members of a particular gene family. The nucleic acid molecules may be  
CC used for isolating a variety of agronomically significant genes  
CC associated with plant growth, quality, yield, and could also serve as  
CC links in metabolic and catabolic pathways. The nucleic acid molecules are  
CC also useful for identifying genes important in initiating and maintaining  
CC seed germination or that may be used to mitigate stresses encountered  
CC during seed germination. The ESTs additionally enable the acquisition of  
CC promoters and cis-regulatory elements which will be useful to express  
CC agronomically significant genes in these tissues and/or other tissues,  
CC and also permits the acquisition of molecular markers useful in breeding  
CC schemes, genetic and molecular mapping, and in cloning of agronomically  
CC significant genes. The nucleic acid molecules are further useful for  
CC detecting the expression level or pattern of a protein or mRNA and for  
CC detecting the presence or quantity of a protein by tissue printing. The  
CC present sequence represents a specifically claimed EST isolated from a  
CC cotton variety Nucotton33B androecium tissue cDNA library (LIB3828). The  
CC sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from the US  
CC patent office at seqdata.uepto.gov/sequence.html?DocID=US20040123340  
XX  
SQ Sequence 482 BP; 137 A; 103 C; 121 G; 121 T; 0 U; 0 Other;  
  
Alignment Scores:  
Pred. No.: 0.000615 Length: 482  
Score: 79.00 Matches: 13  
Percent Similarity: 75.00% Conservative: 5  
Best Local Similarity: 54.17% Mismatches: 6

OS Arabidopsis thaliana.  
 XX WO2003050287-A2.  
 XX 19-JUN-2003.  
 XX 10-DEC-2002; 2002WO-EP013990.  
 XX 10-DEC-2001; 2001US-00014101.  
 XX (SCHM/) SCHMULLING T.  
 XX (WERN/) WERNER T.  
 XX Schmullling T, Werner T;  
 XX WPI; 2003-541577/51.  
 XX P-PSDB; ABR63571.  
 XX Stimulating root growth, enhancing lateral or adventitious root formation  
 PT or altering root geotropism comprises increasing plant cytokinin oxidase  
 PT levels or other protein or nucleic acid that reduces active cytokinins in  
 PT a plant.  
 XX  
 XX Claim 2; Page 153-154; 177pp; English.  
 XX The present invention relates to a method for stimulating root growth or  
 CC enhancing the formation of lateral or adventitious roots or altering root  
 CC geotropism, which comprises increasing in a plant or plant part the level  
 CC of a plant cytokinin oxidase or other protein that reduces the level of  
 CC active cytokinins in a plant or plant part. Cytokinin oxidase protein and  
 CC coding sequences from Arabidopsis thaliana are also provided. The method  
 CC is useful in modifying plant morphological, biochemical and physiological  
 CC properties, such as in modifying the initiation, stimulation or  
 CC enhancement of root growth, adventitious root formation, lateral root  
 CC formation, root geotropism, shoot growth, apical dominance, branching,  
 CC timing of senescence, timing of flowering, flower formation, seed  
 CC development and/or seed yield. The present sequence is a coding sequence  
 CC shown in the invention  
 XX  
 XX Sequence 2782 BP; 826 A; 538 C; 507 G; 911 T; 0 U; 0 Other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 0.00121 Length: 2782  
 Score: 83.00 Matches: 19  
 Percent Similarity: 37.93% Conservative: 3  
 Best Local Similarity: 32.76% Mismatches: 6  
 Query Match: 56.46% Indels: 30  
 DB: 9 Gaps: 1  
 US-10-014-101B-32 (1-28) x ACC85278 (1-2782)  
 QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLys----- 15  
 DB 2357 TCAACTCTGGTGTACTCTCTCTATCCCAACGAAACGAAAGTAATATTACTTTT 2416  
 QY 15 ----- 15  
 DB 2417 TGATTTTGTATTATTGAAAGTATATCCCAATAATGTATGTTAAATGTTTAAACAAGATT 2476  
 QY 16 -----TrpAspAsnArgMetSerAlaMetIleProGluIleAsp 28  
 DB 2477 TATTTTAAATAGATGGAAACACCGCATGTCAACGATGACACCGGACCAAGAT 2530  
 RESULT 28  
 ID ABK28629 standard; cDNA; 1515 BP.  
 XX AC ABK28629;  
 XX 09-APR-2002 (first entry)  
 XX cDNA encoding A. thaliana cytokinin oxidase AtCKX6.  
 XX

KW Arabidopsis thaliana; plant; cytokinin oxidase; transgenic plant;  
 KW root growth; lateral root; adventitious root; root geotropism; herbicide;  
 KW root meristem; shoot meristem; leaf senescence; parthenocarp; gene; ss.  
 XX Arabidopsis thaliana.  
 OS Arabidopsis thaliana.  
 XX WO200196580-A2.  
 XX 20-DEC-2001.  
 PD 18-JUN-2001; 2001WO-EP006833.  
 XX 16-JUN-2000; 2000EP-00870132.  
 XX 27-DEC-2000; 2000US-0258415P.  
 PR 16-MAR-2001; 2001EP-00870053.  
 XX (SCHM/) SCHMULLING T.  
 XX (WERN/) WERNER T.  
 PA Schmullling T, Werner T;  
 PI WPI; 2002-130736/17.  
 XX Polynucleotide encoding novel plant protein having cytokinin oxidase  
 PT activity and the protein useful for stimulating root growth, enhancing  
 PT the formation of lateral or adventitious roots, altering root geotropism.  
 XX Example 1; Page 149; 154pp; English.  
 XX The invention relates to an isolated polynucleotide (I) encoding a novel  
 CC plant protein (II) having cytokinin oxidase activity. (I) is useful for  
 CC production of transgenic plants, plant cells or tissues; for production  
 CC of altered plants, plant cell or tissues; and for effecting the  
 CC expression of (II) where (I) is operably linked to one or more control  
 CC sequences. The methods further comprises regenerating a plant from the  
 CC plant cell. (I) and (II) are useful for stimulating root growth;  
 CC enhancing the formation of lateral or adventitious roots; altering root  
 CC geotropism, leading to an increase in yield; and for screening growth  
 CC promoting chemical of herbicides. (I) is useful for increasing the size  
 CC of the root meristem; increasing root size; increasing the size of the  
 CC shoot meristem; delaying leaf senescence and altering leaf senescence;  
 CC increasing leaf thickness; reducing or increasing the vessel size;  
 CC inducing parthenocarp; improving standability of the seedlings;  
 CC increasing branching and for improving lodging resistance. Antibody (III)  
 CC to (II) is useful for identifying and obtaining proteins interacting with  
 CC (II) comprising a screening assay, preferably a two-hybrid screening  
 CC assay. ABK28606-ABK28633 represent Arabidopsis thaliana cytokinin oxidase  
 CC coding sequences and PCR primers of the invention  
 XX  
 SQ Sequence 1515 BP; 456 A; 327 C; 378 G; 354 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 0.00182 Length: 1515  
 Score: 80.00 Matches: 12  
 Percent Similarity: 80.77% Conservative: 9  
 Best Local Similarity: 46.15% Mismatches: 5  
 Query Match: 54.42% Indels: 0  
 DB: 6 Gaps: 0  
 US-10-014-101B-32 (1-28) x ABK28629 (1-1515)  
 QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20  
 DB 1150 ACAAGCAACGGCCAGTCATCTCTTACCAGTGAACAAATCAAACTGGGACAATCAACA 1209  
 QY 21 SerAlaMetIleProGlu 26  
 DB 1210 TCAGCAGTAACACCGGAG 1227  
 RESULT 29  
 ID ACC85298 standard; cDNA; 1515 BP.  
 XX



```

OS Oryza sativa.
XX
FH Key Location/Qualifiers
FT CDS 64..1761
FT /*tag= a
FT /product= "Rice grain number/cytokinin oxidase (CKX) -
FT related protein - SEQ ID 3"
XX
XX WO2004044200-A1.
XX
XX 27-MAY-2004.
XX
XX 13-NOV-2003; 2003WO-JP014434.
XX
XX 13-NOV-2002; 2002US-0425919P.
XX
XX (HOND ) HONDA MOTOR CO LTD.
XX
XX Ashikari M, Matsuoka M, Lin S, Yamamoto T, Nishimura A;
XX Takashi T;
XX
XX WPI; 2004-420329/39.
XX P-PSDB; ADP03323.
XX
XX Cytokinin oxidase DNA which encodes plant derived protein (CKX) whose
XX functional deletion causes increase in number of grain arrivals in plant,
XX useful for improving plant varieties.
XX
XX Claim 1; SEQ ID NO 2; 89pp; Japanese.
XX
XX The invention relates to a novel cytokinin oxidase (CKX) DNA which
XX encodes the plant-derived protein whose functional deletion causes an
XX increase in the grain number of the plant, including glumous flower,
XX fruit and seed. The molecules of the invention may be useful for
XX producing a transformed plant. Thus, the molecules may be useful for
XX increasing the grain number of a plant and for improving plant varieties.
XX The current sequence is that of the rice grain number/cytokinin oxidase
XX (CKX)-related cDNA (SEQ ID 2) of the invention.
XX
XX Sequence 2302 BP; 428 A; 707 C; 718 G; 449 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 0.000193 Length: 2302
Score: 87.00 Matches: 16
Percent Similarity: 74.07% Conservative: 4
Best Local Similarity: 59.26% Mismatches: 7
Query Match: 59.18% Indels: 0
DB: 12 Gaps: 0
US-10-014-101B-32 (1-28) x ADP03322 (1-2302)
Qy 2 AlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSer 21
Db 1402 GCCATGGGCGCCGCTCATCTACCCCATGACCGCAAGTGGGACAGTAACATGTG 1461
Qy 22 AlaMetIleProGluIleAsp 28
Db 1462 GCAGTGATCACCAGCAGCAGC 1482
RESULT 24
ADA69574
ID ADA69574 standard; DNA; 1587 BP.
XX
XX ADA69574;
XX
XX 20-NOV-2003 (first entry)
XX
XX Rice gene, SEQ ID 2897.
XX
XX Plant; bacterial infection; fungal infection; viral infection; rice;
XX gene; ds.
XX
XX Oryza sativa.
XX

```

```

XX WO2003000898-A1.
XX
XX 03-JAN-2003.
XX
XX 22-JUN-2001; 2001WO-IB001105.
XX
XX 22-JUN-2001; 2001WO-IB001105.
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
XX Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX WPI; 2003-175290/17.
XX
XX Identifying at least one gene involved in plant resistance or response to
XX pathogenic infection for conferring resistance or tolerance to a plant to
XX bacterial, fungal or viral infection by determining or detecting plant
XX gene expression.
XX
XX Claim 6; SEQ ID NO 2897; 899pp; English.
XX
XX The present invention relates to a method (M1) for identifying genes
XX involved in plant resistance or response to pathogenic infection. M1
XX comprises identifying a gene whose expression is significantly altered in
XX the incompatible interaction of plant gene expression relative to
XX expression of the gene in an uninfected plant, in a mutant plant that
XX does not express a gene associated with response to pathogenic infection,
XX or in a corresponding incompatible or compatible interaction. (M1) is
XX useful for conferring resistance to resistance or tolerance to a plant to
XX bacterial, fungal or viral infection. The present sequence was used to
XX illustrate the invention.
XX
XX Sequence 1587 BP; 307 A; 462 C; 458 G; 356 T; 0 U; 4 Other;
XX
Alignment Scores:
Pred. No.: 0.000265 Length: 1587
Score: 85.00 Matches: 14
Percent Similarity: 80.00% Conservative: 6
Best Local Similarity: 56.00% Mismatches: 5
Query Match: 57.82% Indels: 0
DB: 8 Gaps: 0
US-10-014-101B-32 (1-28) x ADA69574 (1-1587)
Qy 4 GlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSerIleMet 23
Db 1255 GGCCTCATCTCATGTACCCCATGAATAAGGACATGTGGGATGACAGGATGACGCGGATG 1314
Qy 24 IleProGluIleAsp 28
Db 1315 ACGCCGAGCAGGAC 1329
RESULT 25
ADA71115
ID ADA71115 standard; DNA; 1566 BP.
XX
XX ADA71115;
XX
XX 20-NOV-2003 (first entry)
XX
XX Rice gene, SEQ ID 4438.
XX
XX Plant; bacterial infection; fungal infection; viral infection; rice;
XX gene; ds.
XX
XX Oryza sativa.
XX
XX WO2003000898-A1.
XX
XX 03-JAN-2003.
XX

```

```
XX DE Rice gene, SEQ ID 3500.
XX KW Plant; bacterial infection; fungal infection; viral infection; rice;
XX KW gene; ds.
XX OS Oryza sativa.
XX PN WO2003000898-A1.
XX PD 03-JAN-2003.
XX PF 22-JUN-2001; 2001WO-IB001105.
XX PR 22-JUN-2001; 2001WO-IB001105.
XX PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
XX PI Katagiri F, Qian S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX DR WPI; 2003-175290/17.
XX PT Identifying at least one gene involved in plant resistance or response to
XX PT pathogenic infection for conferring resistance or tolerance to a plant to
XX PT bacterial, fungal or viral infection by determining or detecting plant
XX PT gene expression.
XX PS Claim 6; SEQ ID NO 3500; 899pp; English.
XX CC The present invention relates to a method (M1) for identifying genes
XX CC involved in plant resistance or response to pathogenic infection. M1
XX CC comprises identifying a gene whose expression is significantly altered in
XX CC the incompatible interaction of plant gene expression relative to
XX CC expression of the gene in an uninfected plant, in a mutant plant that
XX CC does not express a gene associated with response to pathogenic infection,
XX CC or in a corresponding incompatible or compatible interaction. (M1) is
XX CC useful for conferring resistance to resistance or tolerance to a plant to
XX CC bacterial, fungal or viral infection. The present sequence was used to
XX CC illustrate the invention.
XX SQ Sequence 1677 BP; 219 A; 608 C; 582 G; 267 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 0.000128 Length: 1677
Score: 87.00 Matches: 16
Percent Similarity: 74.07% Conservative: 4
Best Local Similarity: 59.26% Mismatches: 7
Query Match: 59.18% Indels: 0
DB: 8 Gaps: 0

US-10-014-101B-32 (1-28) x ADA70177 (1-1677)
QY 2 AlaserGlyLeuAlaLeuLeuTyPrThrAsnArgAsnLysTrpAspAsnArgMetSer 21
Db 1318 GCATGGGCCCGCTCTCTATCTACCCATGACCGACACAGTGGACAGTAACTGTCG 1377
QY 22 AlaMetIleProGluIleAasp 28
Db 1378 GCAGTGATCACCAGCAGCAGCAGC 1398

RESULT 22
ADP03325
ID ADP03325 standard; cDNA; 2282 BP.
XX AC ADP03325;
XX AC ADP03325;
XX DT 12-AUG-2004 (first entry)
XX DE Rice grain number/cytokinin oxidase (CKX)-related cDNA - SEQ ID 5.
XX KW cytokinin oxidase; CKX; grain number; plant; glumous flower; fruit; seed;
XX KW transgenic; rice; ss; gene.

XX OS Oryza sativa.
XX FH Key Location/Qualifiers
XX CDS 48..1739
/*tag= a
/product= "Rice grain number/cytokinin oxidase (CKX)-
related protein - SEQ ID 6"
XX PN WO2004044200-A1.
XX PD 27-MAY-2004.
XX PF 13-NOV-2003; 2003WO-JP014434.
XX PR 13-NOV-2002; 2002US-0425919P.
XX PA (HOND ) HONDA MOTOR CO LTD.
XX PI Ashikari M, Matsuoka M, Lin S, Yamamoto T, Nishimura A;
XX PI Takashi T;
XX DR WPI; 2004-420329/39.
XX DR P-PSDB; ADP03326.
XX PT Cytokinin oxidase DNA which encodes plant derived protein (CKX) whose
XX PT functional deletion causes increase in number of grain arrivals in plant,
XX PT useful for improving plant varieties.
XX PS Disclosure; SEQ ID NO 5; 89pp; Japanese.
XX CC The invention relates to a novel cytokinin oxidase (CKX) DNA which
XX CC encodes the plant-derived protein whose functional deletion causes an
XX CC increase in the grain number of the plant, including glumous flower,
XX CC fruit and seed. The molecules of the invention may be useful for
XX CC producing a transformed plant. Thus, the molecules may be useful for
XX CC increasing the grain number of a plant and for improving plant varieties.
XX CC The current sequence is that of the rice grain number/cytokinin oxidase
XX CC (CKX)-related cDNA (SEQ ID 5) of the invention.
XX SQ Sequence 2282 BP; 423 A; 701 C; 715 G; 443 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0.000191 Length: 2282
Score: 87.00 Matches: 16
Percent Similarity: 74.07% Conservative: 4
Best Local Similarity: 59.26% Mismatches: 7
Query Match: 59.18% Indels: 0
DB: 12 Gaps: 0

US-10-014-101B-32 (1-28) x ADP03325 (1-2282)
QY 2 AlaserGlyLeuAlaLeuLeuTyPrThrAsnArgAsnLysTrpAspAsnArgMetSer 21
Db 1380 GCATGGGCCCGCTCTCTATCTACCCATGACCGACACAGTGGACAGTAACTGTCG 1439
QY 22 AlaMetIleProGluIleAasp 28
Db 1440 GCAGTGATCACCAGCAGCAGCAGC 1460

RESULT 23
ADP03322
ID ADP03322 standard; cDNA; 2302 BP.
XX AC ADP03322;
XX AC ADP03322;
XX DT 12-AUG-2004 (first entry)
XX DE Rice grain number/cytokinin oxidase (CKX)-related cDNA - SEQ ID 2.
XX KW cytokinin oxidase; CKX; grain number; plant; glumous flower; fruit; seed;
XX KW transgenic; rice; ss; gene.
```

```
RESULT 19
ADH61268
ID ADH61268 standard; DNA; 1572 BP.
XX
XX
AC ADH61268;
XX
XX
DT 25-MAR-2004 (first entry)
XX
XX
DE Arabidopsis thaliana cytokinin oxidase (CKX1) DNA #2.
XX
XX
KW Male-sterility; cytokinin oxidase; transgenic plant; seed production;
KW CKX1; mouse-ear cross; gene; ds.
XX
XX
OS Arabidopsis thaliana.
XX
XX
PN US2003163847-A1.
XX
XX
PD 28-AUG-2003.
XX
XX
PF 20-DEC-2002; 2002US-00326184.
XX
XX
PR 20-DEC-2001; 2001US-0343129P.
XX
XX
PA (PHAA ) PHARMACIA CORP.
XX
XX
PI Huang S, Crossland LD, Cheikh N, Morris RO;
XX
XX
DR WPI; 2003-897983/82.
DR GENBANK; AF303979.
XX
XX
PT Producing plants characterized by reversible male-sterility, useful for
PT maintaining male sterility in plants, by transforming a plant cell with a
PT nucleic acid construct containing a polynucleotide encoding a cytokinin
PT oxidase.
XX
XX
PS Claim 42; SEQ ID NO 9; 33pp; English.
XX
XX
CC The invention relates to a method for producing a plant characterised by
CC reversible male-sterility which involves transforming a plant cell with a
CC nucleic acid construct containing a polynucleotide encoding a cytokinin
CC oxidase. The method is useful for producing reversible male-sterility in
CC transgenic plants, or for maintaining male sterility in plants. The
CC method reduces the expense of seed production for existing hybrid plants
CC such as corn, but also makes it possible to produce hybrid varieties of
CC traditionally non-hybrid crops. The method is also useful for introducing
CC economically valuable traits from plants having undesirable production
CC characteristics into plants having desirable characteristics. The present
CC sequence is Arabidopsis thaliana cytokinin oxidase (CKX1) DNA. This
CC sequence is used to illustrate the method of the invention.
XX
XX
SQ Sequence 1572 BP; 461 A; 327 C; 348 G; 436 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.3e-06 Length: 1572
Score: 96.00 Matches: 18
Percent Similarity: 80.77% Conservative: 3
Best Local Similarity: 69.23% Mismatches: 5
Query Match: 65.31% Indels: 0
DB: 10 Gaps: 0

US-10-014-101B-32 (1-28) x ADH61268 (1-1572)
QY 3 SerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22
DQ 1231 AGCGGCTCTGTTCTTTATCTCTATGATCGACACAGTGGATGATCGATCTGCTGCC 1290
QY 23 MetIleProGluLeasp 28
DQ 1291 GCTATACCCGAGGAAGAT 1308

RESULT 20
ADA69773
ID ADA69773 standard; DNA; 1590 BP.
XX
XX
AC ADA69773;
XX
XX
DT 20-NOV-2003 (first entry)
XX
XX
DE Rice gene, SEQ ID 3096.
XX
XX
KW plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.
XX
XX
OS Oryza sativa.
XX
XX
PN WO2003000898-A1.
XX
XX
PD 03-JAN-2003.
XX
XX
PF 22-JUN-2001; 2001WO-IB001105.
XX
XX
PR 22-JUN-2001; 2001WO-IB001105.
XX
XX
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
XX
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX
XX
DR WPI; 2003-175290/17.
XX
XX
PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
XX
PS Claim 6; SEQ ID NO 3096; 899pp; English.
XX
XX
CC The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
XX
SQ Sequence 1590 BP; 356 A; 465 C; 450 G; 309 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 8.05e-05 Length: 1590
Score: 88.00 Matches: 16
Percent Similarity: 73.08% Conservative: 3
Best Local Similarity: 61.54% Mismatches: 7
Query Match: 59.86% Indels: 0
DB: 8 Gaps: 0

US-10-014-101B-32 (1-28) x ADA69773 (1-1590)
QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20
DQ 1222 AGCAACAATGGTCCCATCTACTCTTTACCCAGTGAACAGAACCCAGTGGGACACAGACA 1281
QY 21 SerAlaMetIleProGlu 26
DQ 1282 TCAGTGGTCATACCAGAT 1299
RESULT 21
ADA70177
ID ADA70177 standard; DNA; 1677 BP.
XX
XX
AC ADA70177;
XX
XX
DT 20-NOV-2003 (first entry)
```

XX ABK28626;  
 XX 09-APR-2002 (first entry)  
 XX cDNA encoding A. thaliana cytokinin oxidase AtCKX3.  
 XX Arabidopsis thaliana; plant; cytokinin oxidase; transgenic plant;  
 XX root growth; lateral root; adventitious root; root geotropism; herbicide;  
 XX root meristem; shoot meristem; leaf senescence; parthenocarp; gene; ss.  
 XX Arabidopsis thaliana.  
 XX WO200196580-A2.  
 XX 20-DEC-2001.  
 XX 18-JUN-2001; 2001WO-EP006833.  
 XX 16-JUN-2000; 2000EP-00870132.  
 XX 27-DEC-2000; 2000US-0258415P.  
 XX 16-MAR-2001; 2001EP-00870053.  
 XX (SCHM/) SCHMULLING T.  
 XX (WERN/) WERNER T.  
 XX Schmullling T, Werner T;  
 XX WPI; 2002-130736/17.  
 XX Polynucleotide encoding novel plant protein having cytokinin oxidase  
 XX activity and the protein useful for stimulating root growth, enhancing  
 XX the formation of lateral or adventitious roots, altering root geotropism.  
 XX Claim 3; Page 147; 154pp; English.  
 XX The invention relates to an isolated polynucleotide (I) encoding a novel  
 XX plant protein (II) having cytokinin oxidase activity. (I) is useful for  
 XX production of transgenic plants, plant cells or tissues; for production  
 XX of altered plants, plant cell or tissues; and for effecting the  
 XX expression of (II) where (I) is operably linked to one or more control  
 XX sequences. The methods further comprises regenerating a plant from the  
 XX plant cell. (I) and (II) are useful for stimulating root growth;  
 XX enhancing the formation of lateral or adventitious roots; altering root  
 XX geotropism, leading to an increase in yield; and for screening growth  
 XX promoting chemical of herbicides. (I) is useful for increasing the size  
 XX of the root meristem; increasing root size; increasing the size of the  
 XX shoot meristem; delaying leaf senescence and altering leaf senescence;  
 XX increasing leaf thickness; reducing or increasing the vessel size;  
 XX inducing parthenocarp; improving standability of the seedlings;  
 XX increasing branching and for improving lodging resistance. Antibody (III)  
 XX to (II) is useful for identifying and obtaining proteins interacting with  
 XX (II) comprising a screening assay, preferably a two-hybrid screening  
 XX assay. ABK28606-ABK28633 represent Arabidopsis thaliana cytokinin oxidase  
 XX coding sequences and PCR primers of the invention  
 XX SQ Sequence 1572 BP; 461 A; 327 C; 348 G; 436 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 3.3e-06 Length: 1572  
 Score: 96.00 Matches: 18  
 Percent Similarity: 80.77% Conservative: 3  
 Best Local Similarity: 69.23% Mismatches: 5  
 Query Match: 65.31% Indels: 0  
 DB: Gaps: 0  
 US-10-014-101B-32 (1-28) x ABK28626 (1-1572)  
 QY 3 SerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAsnArgMetSerAla 22  
 DB 1231 AGCGGTCTGTTCTTTATCTATGATCGCAACAGTGAATGATCGGATGTCGTC 1290  
 QY 23 MetIleProGluIleAsp 28

Db 1291 GCTATACCGAGGAGAT 1308  
 RESULT 18  
 ACC85295  
 ID ACC85295 standard; cDNA; 1572 BP.  
 XX ACC85295;  
 XX 18-SEP-2003 (first entry)  
 XX Arabidopsis cytokinin oxidase-like protein 3 cDNA.  
 XX Maize; root growth; root geotropism; cytokinin oxidase; seed size;  
 XX embryo size; cotyledon size; transgenic plant; herbicide; gene; ss.  
 XX Arabidopsis thaliana.  
 XX WO2003050287-A2.  
 XX 19-JUN-2003.  
 XX 10-DEC-2002; 2002WO-EP013990.  
 XX 10-DEC-2001; 2001US-00014101.  
 XX (SCHM/) SCHMULLING T.  
 XX (WERN/) WERNER T.  
 XX Schmullling T, Werner T;  
 XX WPI; 2003-541577/51.  
 XX Stimulating root growth, enhancing lateral or adventitious root formation  
 XX or altering root geotropism comprises increasing plant cytokinin oxidase  
 XX levels or other protein or nucleic acid that reduces active cytokinins in  
 XX a plant.  
 XX Claim 3; Page 168-169; 177pp; English.  
 XX The present invention relates to a method for stimulating root growth or  
 XX enhancing the formation of lateral or adventitious roots or altering root  
 XX geotropism, which comprises increasing in a plant or plant part the level  
 XX of a plant cytokinin oxidase or other protein that reduces the level of  
 XX active cytokinins in a plant or plant part. Cytokinin oxidase protein and  
 XX coding sequences from Arabidopsis thaliana are also provided. The method  
 XX is useful in modifying plant morphological, biochemical and physiological  
 XX properties, such as in modifying the initiation, stimulation or  
 XX enhancement of root growth, adventitious root formation, lateral root  
 XX formation, root geotropism, shoot growth, apical dominance, branching,  
 XX timing of senescence, timing of flowering, flower formation, seed  
 XX development and/or seed yield. The present sequence is a coding sequence  
 XX shown in the invention  
 XX SQ Sequence 1572 BP; 461 A; 327 C; 348 G; 436 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 3.3e-06 Length: 1572  
 Score: 96.00 Matches: 18  
 Percent Similarity: 80.77% Conservative: 3  
 Best Local Similarity: 69.23% Mismatches: 5  
 Query Match: 65.31% Indels: 0  
 DB: Gaps: 0  
 US-10-014-101B-32 (1-28) x ACC85295 (1-1572)  
 QY 3 SerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAsnArgMetSerAla 22  
 DB 1231 AGCGGTCTGTTCTTTATCTATGATCGCAACAGTGAATGATCGGATGTCGTC 1290  
 QY 23 MetIleProGluIleAsp 28  
 Db 1291 GCTATACCGAGGAGAT 1308

XX PD 20-DEC-2001.  
 XX XX  
 XX PF 18-JUN-2001; 2001WO-EP006833.  
 XX XX  
 XX PR 16-JUN-2000; 2000EP-00870132.  
 XX PR 27-DEC-2000; 2000US-0258415P.  
 XX PR 16-MAR-2001; 2001EP-00870053.  
 XX XX  
 XX PA (SCHM/) SCHMULLING T.  
 XX PA (WERN/) WERNER T.  
 XX PI Schmullling T, Werner T;  
 XX PD P-PSDB; AAU81968.  
 XX XX  
 XX WPI; 2002-130736/17.  
 XX DR P-PSDB; AAU81968.  
 XX XX  
 XX PT Polynucleotide encoding novel plant protein having cytokinin oxidase  
 XX PT activity and the protein useful for stimulating root growth, enhancing  
 XX PT the formation of lateral or adventitious roots, altering root geotropism.  
 XX PS Claim 3; Page 128-129; 154pp; English.  
 XX XX  
 XX CC The invention relates to an isolated polynucleotide (I) encoding a novel  
 XX CC plant protein (II) having cytokinin oxidase activity. (I) is useful for  
 XX CC production of transgenic plants, plant cells or tissues; for production  
 XX CC of altered plants, plant cell or tissues; and for effecting the  
 XX CC expression of (II) where (I) is operably linked to one or more control  
 XX CC sequences. The methods further comprises regenerating a plant from the  
 XX CC plant cell. (I) and (II) are useful for stimulating root growth;  
 XX CC enhancing the formation of lateral or adventitious roots; altering root  
 XX CC geotropism, leading to an increase in yield; and for screening growth  
 XX CC promoting chemical of herbicides. (I) is useful for increasing the size  
 XX CC of the root meristem; increasing root size; increasing the size of the  
 XX CC shoot meristem; delaying leaf senescence and altering leaf senescence;  
 XX CC increasing leaf thickness; reducing or increasing the vessel size;  
 XX CC inducing parthenocarp; improving standability of the seedlings;  
 XX CC increasing branching and for improving lodging resistance. Antibody (III)  
 XX CC to (II) is useful for identifying and obtaining proteins interacting with  
 XX CC (II) comprising a screening assay, preferably a two-hybrid screening  
 XX CC assay. ABK28606-ABK28633 represent Arabidopsis thaliana cytokinin oxidase  
 XX CC coding sequences and PCR primers of the invention  
 XX SQ Sequence 2991 BP; 920 A; 522 C; 521 G; 1028 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 6.22e-06 Length: 2991  
 Score: 96.50 Matches: 27  
 Percent Similarity: 48.21% Conservative: 0  
 Best Local Similarity: 48.21% Mismatches: 1  
 Query Match: 65.65% Indels: 29  
 DB: 6 Gaps: 1  
 US-10-014-101B-32 (1-28) x ABK28607 (1-2991)  
 QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLys----- 15  
 Db 2573 TCAGCTTCGGGACTCGCTCTCTCTATCCAAACCGGAATAA-GTACATACTTCTCT 2631  
 QY 15 ----- 15  
 Db 2632 CATTCATATTTATCTTCAAGAACCAAGTAATAAATTTCTATGAAGTATTATGCTGTT 2691  
 QY 16 -----TrpAspAsnArgMetSerAlaMetIleProGluIleAsp 28  
 Db 2692 ATTGTTAGATGGGACAATCGTATGTCGGCGATGATACAGAGATCGAT 2739  
 RESULT 16  
 ID ACC85276 standard; DNA; 2991 BP.  
 XX AC ACC85276;  
 XX XX

DT 18-SEP-2003 (first entry)  
 XX Arabidopsis cytokinin oxidase-like protein 2 gene.  
 XX XX  
 XX KW Maize; root growth; root geotropism; cytokinin oxidase; seed size;  
 XX KW embryo size; cotyledon size; transgenic plant; herbicide; gene; ds.  
 XX OS Arabidopsis thaliana.  
 XX XX  
 XX PN WO2003050287-A2.  
 XX XX  
 XX PD 19-JUN-2003.  
 XX XX  
 XX PF 10-DEC-2002; 2002WO-EP013990.  
 XX XX  
 XX PR 10-DEC-2001; 2001US-00014101.  
 XX XX  
 XX PA (SCHM/) SCHMULLING T.  
 XX PA (WERN/) WERNER T.  
 XX PI Schmullling T, Werner T;  
 XX PD P-PSDB; ABR63569.  
 XX DR WPI; 2003-541577/51.  
 XX DR P-PSDB; ABR63569.  
 XX XX  
 XX PT Stimulating root growth, enhancing lateral or adventitious root formation  
 XX PT or altering root geotropism comprises increasing plant cytokinin oxidase  
 XX PT levels or other protein or nucleic acid that reduces active cytokinins in  
 XX PT a plant.  
 XX PS Claim 3; Page 145-146; 177pp; English.  
 XX XX  
 XX CC The present invention relates to a method for stimulating root growth or  
 XX CC enhancing the formation of lateral or adventitious roots or altering root  
 XX CC geotropism, which comprises increasing in a plant or plant part the level  
 XX CC of a plant cytokinin oxidase or other protein that reduces the level of  
 XX CC active cytokinins in a plant or plant part. Cytokinin oxidase protein and  
 XX CC coding sequences from Arabidopsis thaliana are also provided. The method  
 XX CC is useful in modifying plant morphological, biochemical and physiological  
 XX CC properties, such as in modifying the initiation, stimulation or  
 XX CC enhancement of root growth, adventitious root formation, lateral root  
 XX CC formation, root geotropism, shoot growth, apical dominance, branching,  
 XX CC timing of senescence, timing of flowering, flower formation, seed  
 XX CC development and/or seed yield. The present sequence is a coding sequence  
 XX CC shown in the invention  
 XX SQ Sequence 2991 BP; 920 A; 522 C; 521 G; 1028 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 6.22e-06 Length: 2991  
 Score: 96.50 Matches: 27  
 Percent Similarity: 48.21% Conservative: 0  
 Best Local Similarity: 48.21% Mismatches: 1  
 Query Match: 65.65% Indels: 29  
 DB: 9 Gaps: 1  
 US-10-014-101B-32 (1-28) x ACC85276 (1-2991)  
 QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLys----- 15  
 Db 2573 TCAGCTTCGGGACTCGCTCTCTCTATCCAAACCGGAATAA-GTACATACTTCTCT 2631  
 QY 15 ----- 15  
 Db 2632 CATTCATATTTATCTTCAAGAACCAAGTAATAAATTTCTATGAAGTATTATGCTGTT 2691  
 QY 16 -----TrpAspAsnArgMetSerAlaMetIleProGluIleAsp 28  
 Db 2692 ATTGTTAGATGGGACAATCGTATGTCGGCGATGATACAGAGATCGAT 2739  
 RESULT 17  
 ID ABK28626  
 ID ABK28626 standard; cDNA; 1572 BP.

KW CKX1; mouse-ear cross; gene; ds.  
 XX Arabidopsis thaliana.  
 XX US2003163847-A1.  
 XX 28-AUG-2003.  
 XX 20-DEC-2002; 2002US-00326184.  
 XX 20-DEC-2001; 2001US-0343129P.  
 XX (PHAA ) PHARMACIA CORP.  
 XX Huang S, Crossland LD, Cheikh N, Morris RO;  
 PI WPI; 2003-897983/82.  
 XX GENBANK; AF303980.  
 XX Producing plants characterized by reversible male-sterility, useful for  
 PT maintaining male sterility in plants, by transforming a plant cell with a  
 PT nucleic acid construct containing a polynucleotide encoding a cytokinin  
 PT oxidase.  
 XX Claim 43; SEQ ID NO 10; 33pp; English.  
 XX The invention relates to a method for producing a plant characterised by  
 CC reversible male-sterility which involves transforming a plant cell with a  
 CC nucleic acid construct containing a polynucleotide encoding a cytokinin  
 CC oxidase. The method is useful for producing reversible male-sterility in  
 CC transgenic plants, or for maintaining male sterility in plants. The  
 CC method reduces the expense of seed production for existing hybrid plants  
 CC such as corn, but also makes it possible to produce hybrid varieties of  
 CC traditionally valuable traits from plants having undesirable production  
 CC characteristics into plants having desirable characteristics. The present  
 CC sequence is Arabidopsis thaliana cytokinin oxidase (CKX1) DNA. This  
 CC sequence is used to illustrate the method of the invention.  
 XX SQ Sequence 1575 BP; 418 A; 367 C; 349 G; 441 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 2 81e-08 Length: 1575  
 Score: 108.00 Matches: 19  
 Percent Similarity: 78.57% Conservative: 3  
 Best Local Similarity: 67.86% Mismatches: 6  
 Query Match: 73.47% Indels: 0  
 DB: 10 Gaps: 0  
 US-10-014-101B-32 (1-28) x ADH61269 (1-1575)  
 QY 1 SerAlaSerGlyLeuAlaLeuLeuTyrProThrAsnArgAsnLysTrpAspAsnArgMet 20  
 Db 1240 TCAACTTCGTGTTACTCTCTCTATCCACAAACCGAAACAAATGGAAACACCGCATG 1299  
 QY 21 SerAlaMetIleProGluIleAsp 28  
 Db 1300 TCAACGATGACACCGACGAAGAT 1323  
 RESULT 14  
 ADH61279  
 ID ADH61279 standard; DNA; 1655 BP.  
 XX AC ADH61279;  
 XX 25-MAR-2004 (first entry)  
 XX Arabidopsis thaliana cytokinin oxidase (CKX1) DNA #6.  
 XX Male-sterility; cytokinin oxidase; transgenic plant; seed production;  
 KW CKX1; mouse-ear cross; gene; ds.  
 XX Arabidopsis thaliana.

XX US2003163847-A1.  
 XX 28-AUG-2003.  
 XX 20-DEC-2002; 2002US-00326184.  
 XX 20-DEC-2001; 2001US-0343129P.  
 XX (PHAA ) PHARMACIA CORP.  
 XX Huang S, Crossland LD, Cheikh N, Morris RO;  
 PI WPI; 2003-897983/82.  
 XX GENBANK; BT000179.  
 XX Producing plants characterized by reversible male-sterility, useful for  
 PT maintaining male sterility in plants, by transforming a plant cell with a  
 PT nucleic acid construct containing a polynucleotide encoding a cytokinin  
 PT oxidase.  
 XX Disclosure; SEQ ID NO 20; 33pp; English.  
 XX The invention relates to a method for producing a plant characterised by  
 CC reversible male-sterility which involves transforming a plant cell with a  
 CC nucleic acid construct containing a polynucleotide encoding a cytokinin  
 CC oxidase. The method is useful for producing reversible male-sterility in  
 CC transgenic plants, or for maintaining male sterility in plants. The  
 CC method reduces the expense of seed production for existing hybrid plants  
 CC such as corn, but also makes it possible to produce hybrid varieties of  
 CC traditionally valuable traits from plants having undesirable production  
 CC characteristics into plants having desirable characteristics. The present  
 CC sequence is Arabidopsis thaliana cytokinin oxidase (CKX1) DNA. This  
 CC sequence is used to illustrate the method of the invention.  
 XX SQ Sequence 1655 BP; 444 A; 380 C; 358 G; 473 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 2 99e-08 Length: 1655  
 Score: 108.00 Matches: 19  
 Percent Similarity: 78.57% Conservative: 3  
 Best Local Similarity: 67.86% Mismatches: 6  
 Query Match: 73.47% Indels: 0  
 DB: 10 Gaps: 0  
 US-10-014-101B-32 (1-28) x ADH61279 (1-1655)  
 QY 1 SerAlaSerGlyLeuAlaLeuLeuTyrProThrAsnArgAsnLysTrpAspAsnArgMet 20  
 Db 1240 TCAACTTCGTGTTACTCTCTCTATCCACAAACCGAAACAAATGGAAACACCGCATG 1299  
 QY 21 SerAlaMetIleProGluIleAsp 28  
 Db 1300 TCAACGATGACACCGACGAAGAT 1323  
 RESULT 15  
 ABK28607  
 ID ABK28607 standard; DNA; 2991 BP.  
 XX AC ABK28607;  
 XX 09-APR-2002 (first entry)  
 XX DNA encoding A. thaliana cytokinin oxidase AtCKX2.  
 XX Arabidopsis thaliana; plant; cytokinin oxidase; transgenic plant;  
 KW root growth; lateral root; adventitious root; root geotropism; herbicide;  
 KW root meristem; shoot meristem; leaf senescence; parthenocarpy; gene; ds.  
 XX Arabidopsis thaliana.  
 XX WO200196580-A2.







PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155569P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 21-OCT-1999; 99US-0160880P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161559P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

## Alignment Scores:

Pred. No.: 4,38e-12 Length: 1548  
Score: 130.00 Matches: 28  
Percent Similarity: 66.67% Conservative: 0  
Best Local Similarity: 66.67% Mismatches: 0  
Query Match: 88.44% Indels: 14  
DB: 3 Gaps: 1

US-10-014-101b-32 (1-28) x AAC43214 (1-1548)

QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsn----- 14  
DB 1171 TCAGCTTCGGAGCTCGTCTCTCTATCCAAACCGGAATAGTACATATCTCTCTC 1230  
QY 15 -----LysTrpAspAsnArgMetSerAlaMetIleProGlu 26  
DB 1231 ATTCAATTATTATCTTCAGNACCAAAATGGCAATCGTATGCGGATGATACCAGAG 1290  
QY 27 IleAsp 28  
DB 1291 ATCGAT 1296

## RESULT 8

AAC42983  
ID AAC42983 standard; DNA; 1575 BP.  
XX  
AC AAC42983;  
XX  
DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 37588.  
DE Hybridisation assay; Genetic mapping; Gene expression control;  
XX protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.  
XX Arabidopsis thaliana.  
OS Arabidopsis thaliana.  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-00301439.  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0128845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132484P.  
PR 06-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0132863P.  
PR 14-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 21-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.

XX Arabidopsis thaliana.  
XX EP1033405-A2.  
XX 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-00301439.  
XX 25-FEB-1999; 99US-0121825P.  
XX 05-MAR-1999; 99US-0123180P.  
XX 03-MAR-1999; 99US-0123548P.  
XX 23-MAR-1999; 99US-0125788P.  
XX 25-MAR-1999; 99US-0126264P.  
XX 01-APR-1999; 99US-0126785P.  
XX 06-APR-1999; 99US-0127462P.  
XX 08-APR-1999; 99US-0128234P.  
XX 16-APR-1999; 99US-0128714P.  
XX 19-APR-1999; 99US-0129845P.  
XX 21-APR-1999; 99US-0130077P.  
XX 23-APR-1999; 99US-0130449P.  
XX 28-APR-1999; 99US-0130891P.  
XX 30-APR-1999; 99US-0131449P.  
XX 04-MAY-1999; 99US-0132407P.  
XX 05-MAY-1999; 99US-0132484P.  
XX 06-MAY-1999; 99US-0132485P.  
XX 07-MAY-1999; 99US-0132487P.  
XX 11-MAY-1999; 99US-0132863P.  
XX 14-MAY-1999; 99US-0134218P.  
XX 14-MAY-1999; 99US-0134219P.  
XX 18-MAY-1999; 99US-0134370P.  
XX 19-MAY-1999; 99US-0134768P.  
XX 20-MAY-1999; 99US-0134941P.  
XX 21-MAY-1999; 99US-0135124P.  
XX 24-MAY-1999; 99US-0135353P.  
XX 25-MAY-1999; 99US-0135629P.  
XX 27-MAY-1999; 99US-0136021P.  
XX 28-MAY-1999; 99US-0136332P.  
XX 01-JUN-1999; 99US-0136782P.  
XX 03-JUN-1999; 99US-0137222P.  
XX 04-JUN-1999; 99US-0137502P.  
XX 07-JUN-1999; 99US-0137724P.  
XX 08-JUN-1999; 99US-0138094P.  
XX 10-JUN-1999; 99US-0138540P.  
XX 10-JUN-1999; 99US-0138847P.  
XX 14-JUN-1999; 99US-0139119P.  
XX 16-JUN-1999; 99US-0139452P.  
XX 16-JUN-1999; 99US-0139453P.  
XX 17-JUN-1999; 99US-0139452P.  
XX 18-JUN-1999; 99US-0139454P.  
XX 18-JUN-1999; 99US-0139455P.  
XX 18-JUN-1999; 99US-0139456P.  
XX 18-JUN-1999; 99US-0139457P.  
XX 18-JUN-1999; 99US-0139458P.  
XX 18-JUN-1999; 99US-0139459P.  
XX 18-JUN-1999; 99US-0139460P.  
XX 18-JUN-1999; 99US-0139461P.  
XX 18-JUN-1999; 99US-0139462P.  
XX 18-JUN-1999; 99US-0139463P.  
XX 18-JUN-1999; 99US-0139750P.  
XX 21-JUN-1999; 99US-0139763P.  
XX 21-JUN-1999; 99US-0139817P.  
XX 22-JUN-1999; 99US-0139899P.  
XX 23-JUN-1999; 99US-0140353P.  
XX 24-JUN-1999; 99US-0140354P.  
XX 24-JUN-1999; 99US-0140695P.  
XX 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 08-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.

```
XX 20-DEC-2001; 2001US-0343129P.
XX (PHAA ) PHARMACIA CORP.
XX Huang S, Crossland LD, Cheikh N, Morris RO;
XX WPI; 2003-897983/82.
XX GENBANK; AF303978.
XX Producing plants characterized by reversible male-sterility, useful for
XX maintaining male sterility in plants, by transforming a plant cell with a
XX nucleic acid construct containing a polynucleotide encoding a cytokinin
XX oxidase.
XX Claim 41; SEQ ID NO 8; 33pp; English.
XX The invention relates to a method for producing a plant characterised by
XX reversible male-sterility which involves transforming a plant cell with a
XX nucleic acid construct containing a polynucleotide encoding a cytokinin
XX oxidase. The method is useful for producing reversible male-sterility in
XX transgenic plants, or for maintaining male sterility in plants. The
XX method reduces the expense of seed production for existing hybrid plants
XX such as corn, but also makes it possible to produce hybrid varieties of
XX traditionally non-hybrid crops. The method is also useful for introducing
XX economically valuable traits from plants having undesirable production
XX characteristics into plants having desirable characteristics. The present
XX sequence is Arabidopsis thaliana cytokinin oxidase (CKX1) DNA. This
XX sequence is used to illustrate the method of the invention.
XX Sequence 1506 BP; 431 A; 329 C; 323 G; 423 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 4,91e-15 Length: 1506
Score: 147.00 Matches: 28
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0
US-10-014-101B-32 (1-28) x ADH61267 (1-1506)
Qy 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20
Db 1171 TCAGCTTCGGGACTCGCTCTCTCTATCCAAACCGGAATAATGGGACATCGTATG 1230
Qy 21 SerAlaMetIleProGluIleAsp 28
Db 1231 TCGGCGATGATACAGAGATCGAT 1254
RESULT 6
AD006498
ID AD006498 standard; DNA; 1506 BP.
AC AD006498;
XX 29-JUL-2004 (first entry)
XX A thaliana cytokinin oxidase AtCKX2 coding sequence.
XX bioremediation; AtCKX2; cytokinin availability; contaminant;
XX metal deficiency; nutrition; ds; gene; cytokinin oxidase.
XX Arabidopsis thaliana.
XX Key Location/Qualifiers
XX CDS 1..1506
XX /*tag= a
XX /product= "AtCKX2"
XX WO2004038027-A1.
XX 06-MAY-2004.
XX
```

```
XX 24-OCT-2003; 2003WO-EP012051.
XX 24-OCT-2002; 2002EP-00079481.
XX (CROP-) CROPDESIGN NV.
XX Van Camp W;
XX WPI; 2004-375913/35.
XX P-PSDB; ADO06499.
XX Method for bioremediation, useful for removing contaminants or metals, by
XX decreasing cytokinin availability in plants, and cultivating plant on
XX substrate comprising one or more contaminants.
XX Claim 8; Page 55; 61pp; English.
XX The present invention relates to a method for bioremediation, which
XX involves decreasing cytokinin availability in a plant relative to a
XX corresponding wild type plant, and cultivating the plant on a substrate
XX comprising one or more contaminants, or cultivating a plant having
XX lowered availability of cytokinin relative to corresponding wild type
XX plants, on a substrate, which is to be treated. The method is useful for
XX bioremediation, for concentrating contaminants in a plant, where the
XX plant has a higher concentration of contaminants compared to a
XX corresponding wild type plant. The metal contaminants include aluminum,
XX americium, antimony, arsenic, barium, beryllium, bismuth, cadmium,
XX caesium, cerium, chromium, copper, gallium, germanium, gold, indium,
XX iridium, iron, lead, manganese, mercury, molybdenum, neptunium, osmium,
XX palladium, platinum, plutonium, radium, rhenium, rhodium, rubidium,
XX ruthenium, scandium, selenium, silver, strontium, technetium, tellurium,
XX thallium, tin, tungsten, uranium, vanadium or yttrium, preferably
XX cadmium. The plant obtained by the method is useful in bioremediation.
XX Transgenic plants with a lower availability of cytokinin are useful in
XX the manufacture of a medicament for treatment of disorders arising from
XX metal deficiencies, and as a medicament for improving animal or human
XX nutrition. The present sequence is a coding sequence shown in the
XX exemplification of the invention.
XX Sequence 1506 BP; 431 A; 328 C; 324 G; 423 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 4,91e-15 Length: 1506
Score: 147.00 Matches: 28
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0
US-10-014-101B-32 (1-28) x ADO06498 (1-1506)
Qy 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20
Db 1171 TCAGCTTCGGGACTCGCTCTCTCTATCCAAACCGGAATAATGGGACATCGTATG 1230
Qy 21 SerAlaMetIleProGluIleAsp 28
Db 1231 TCGGCGATGATACAGAGATCGAT 1254
RESULT 7
AAC43214
ID AAC43214 standard; DNA; 1548 BP.
XX AAC43214;
XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 38443.
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway; metabolic pathway;
XX promoter; termination sequence; ss.
XX
```

```
PR 16-MAR-2001; 2001EP-00870053.
XX (SCHM/) SCHMULLING T.
PA (WERN/) WERNER T.
XX
XX Schmullling T, Werner T;
XX
XX WPI; 2002-130736/17.
XX
XX Polynucleotide encoding novel plant protein having cytokinin oxidase
PT activity and the protein useful for stimulating root growth, enhancing
PT the formation of lateral or adventitious roots, altering root geotropism.
XX
XX Claim 2; Page 146-147; 154pp; English.
XX
XX The invention relates to an isolated polynucleotide (I) encoding a novel
CC plant protein (II) having cytokinin oxidase activity. (I) is useful for
CC production of transgenic plants, plant cells or tissues; for production
CC of altered plants, plant cell or tissues; and for effecting the
CC expression of (II) where (I) is operably linked to one or more control
CC sequences. The methods further comprises regenerating a plant from the
CC plant cell. (I) and (II) are useful for stimulating root growth;
CC enhancing the formation of lateral or adventitious roots; altering root
CC geotropism, leading to an increase in yield; and for screening growth
CC promoting chemical of herbicides. (I) is useful for increasing the size
CC of the root meristem; increasing root size; increasing the size of the
CC shoot meristem; delaying leaf senescence and altering leaf senescence;
CC increasing leaf thickness; reducing or increasing the vessel size;
CC inducing parthenocarp; improving standability of the seedlings;
CC increasing branching and for improving lodging resistance. Antibody (III)
CC to (II) is useful for identifying and obtaining proteins interacting with
CC (II) comprising a screening assay, preferably a two-hybrid screening
CC assay. ABK28606-ABK28633 represent Arabidopsis thaliana cytokinin oxidase
CC coding sequences and PCR primers of the invention
XX
XX Sequence 1506 BP; 431 A; 328 C; 324 G; 423 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 4.91e-15 Length: 1506
Score: 147.00 Matches: 28
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-014-101B-32 (1-28) x ABK28625 (1-1506)
QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20
Db 1171 TCAGCTTCGGGACTCGCTCTCTCTATCCAAACCGGAAATGAGGACATCGTATG 1230
QY 21 SerAlaMetIleProGluIleAsp 28
Db 1231 TCGGCGATGATACAGAGATCGAT 1254
RESULT 4
ID ACC85294 standard; cDNA; 1506 BP.
XX
XX ACC85294;
AC ACC85294;
XX
XX 18-SEP-2003 (first entry)
DT
DE Arabidopsis cytokinin oxidase-like protein 2 cDNA.
XX
XX Maize; root growth; root geotropism; cytokinin oxidase; seed size;
XX embryo size; cotyledon size; transgenic plant; herbicide; gene; ss.
XX Arabidopsis thaliana.
XX
XX WO2003050287-A2.
PN
XX 19-JUN-2003.
PD
10-DEC-2002; 2002WO-EF013990.
XX
XX 10-DEC-2001; 2001US-00014101.
XX
XX (SCHM/) SCHMULLING T.
PA (WERN/) WERNER T.
XX
XX Schmullling T, Werner T;
XX
XX WPI; 2003-541577/51.
XX
XX Stimulating root growth, enhancing lateral or adventitious root formation
PT or altering root geotropism comprises increasing plant cytokinin oxidase
PT levels or other protein or nucleic acid that reduces active cytokinins in
PT a plant.
XX
XX Claim 3; Page 167-168; 177pp; English.
XX
XX The present invention relates to a method for stimulating root growth or
CC enhancing the formation of lateral or adventitious roots or altering root
CC geotropism, which comprises increasing in a plant or plant part the level
CC of a plant cytokinin oxidase or other protein that reduces the level of
CC active cytokinins in a plant or plant part. Cytokinin oxidase protein and
CC coding sequences from Arabidopsis thaliana are also provided. The method
CC is useful in modifying plant morphological, biochemical and physiological
CC properties, such as in modifying the initiation, stimulation or
CC enhancement of root growth, adventitious root formation, lateral root
CC formation, root geotropism, shoot growth, apical dominance, branching,
CC timing of senescence, timing of flowering, flower formation, seed
CC development and/or seed yield. The present sequence is a coding sequence
CC shown in the invention
XX
XX Sequence 1506 BP; 431 A; 328 C; 324 G; 423 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 4.91e-15 Length: 1506
Score: 147.00 Matches: 28
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-10-014-101B-32 (1-28) x ACC85294 (1-1506)
QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20
Db 1171 TCAGCTTCGGGACTCGCTCTCTCTATCCAAACCGGAAATGAGGACATCGTATG 1230
QY 21 SerAlaMetIleProGluIleAsp 28
Db 1231 TCGGCGATGATACAGAGATCGAT 1254
RESULT 5
ID ADH61267 standard; DNA; 1506 BP.
XX
XX ADH61267;
AC ADH61267;
XX
XX 25-MAR-2004 (first entry)
DT
DE Arabidopsis thaliana cytokinin oxidase (CKX1) DNA #1.
XX
XX Male-sterility; cytokinin oxidase; transgenic plant; seed production;
XX CKX1; mouse-ear cross; gene; ds.
XX Arabidopsis thaliana.
XX
XX US2003163847-A1.
PN
XX 28-AUG-2003.
PD
20-DEC-2002; 2002US-00326184.
PF
```

PI Schmulling T, Werner T;  
 XX WPI; 2002-130736/17.  
 DR P-PSDB; AAU81973.  
 XX  
 XX Polynucleotide encoding novel plant protein having cytokinin oxidase  
 PT activity and the protein useful for stimulating root growth, enhancing  
 PT the formation of lateral or adventitious roots, altering root geotropism.  
 PS  
 PS Claim 3; Page 149; 154pp; English.  
 XX  
 XX The invention relates to an isolated polynucleotide (I) encoding a novel  
 CC plant protein (II) having cytokinin oxidase activity. (I) is useful for  
 CC production of transgenic plants, plant cells or tissues; for production  
 CC of altered plants, plant cell or tissues; and for effecting the  
 CC expression of (II) where (I) is operably linked to one or more control  
 CC sequences. The methods further comprises regenerating a plant from the  
 CC plant cell. (I) and (II) are useful for stimulating root growth;  
 CC enhancing the formation of lateral or adventitious roots; altering root  
 CC geotropism, leading to an increase in yield; and for screening growth  
 CC promoting chemical of herbicides. (I) is useful for increasing the size  
 CC of the root meristem; increasing root size; increasing the size of the  
 CC shoot meristem; delaying leaf senescence and altering leaf senescence;  
 CC increasing leaf thickness; reducing or increasing the vessel size;  
 CC inducing parthenocary; improving standability of the seedlings;  
 CC increasing branching and for improving lodging resistance. Antibody (III)  
 CC to (II) is useful for identifying and obtaining proteins interacting with  
 CC (II) comprising a screening assay, preferably a two-hybrid screening  
 CC assay. ABK28606-ABK28633 represent Arabidopsis thaliana cytokinin oxidase  
 CC coding sequences and PCR primers of the invention  
 XX  
 SQ Sequence 84 BP; 24 A; 21 C; 19 G; 20 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 1.17e-16 Length: 84  
 Score: 147.00 Matches: 28  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-10-014-101B-32 (1-28) x ABK28630 (1-84)

QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20  
 DB 1 TCAGCTTCGGGACTCGCTCTCTCTATCCAAACCGGAATAAATGGGACATCGTATG 60  
 QY 21 SerAlaMetIleProGluIleAsp 28  
 DB 61 TCGGCGATGATACACAGATCGAT 84

RESULT 2  
 ACC85299  
 ID ACC85299 standard; cDNA; 84 BP.

AC. ACC85299;

XX 19-SEP-2003 (first entry)

DE Arabidopsis cytokinin oxidase-like protein 2 cDNA fragment #1.

XX Maize; root growth; root geotropism; cytokinin oxidase; seed size;  
 KW embryo size; cotyledon size; transgenic plant; herbicide; gene; ss.

XX Arabidopsis thaliana.

XX WO2003050287-A2.

PN 19-JUN-2003.

XX 10-DEC-2002; 2002WO-EP013990.

PF 10-DEC-2001; 2001US-00014101.

PR

XX (SCHM/) SCHMULLING T.  
 PA (WERN/) WERNER T.  
 XX

PI Schmulling T, Werner T;

XX WPI; 2003-541577/51.

DR Stimulating root growth, enhancing lateral or adventitious root formation  
 XX or altering root geotropism comprises increasing plant cytokinin oxidase  
 PT levels or other protein or nucleic acid that reduces active cytokinins in  
 PT a plant.

XX Claim 3; Page 172; 177pp; English.

XX The present invention relates to a method for stimulating root growth or  
 PS enhancing the formation of lateral or adventitious roots or altering root  
 CC geotropism, which comprises increasing in a plant or plant part the level  
 CC of a plant cytokinin oxidase or other protein that reduces the level of  
 CC active cytokinins in a plant or plant part. Cytokinin oxidase protein and  
 CC coding sequences from Arabidopsis thaliana are also provided. The method  
 CC is useful in modifying plant morphological, biochemical and physiological  
 CC properties, such as in modifying the initiation, stimulation or  
 CC enhancement of root growth, adventitious root formation, lateral root  
 CC formation, root geotropism, shoot growth, apical dominance, branching,  
 CC timing of senescence, timing of flowering, flower formation, seed  
 CC development and/or seed yield. The present sequence is a coding sequence  
 CC shown in the invention

XX Sequence 84 BP; 24 A; 21 C; 19 G; 20 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.17e-16 Length: 84  
 Score: 147.00 Matches: 28  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-10-014-101B-32 (1-28) x ACC85299 (1-84)

QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20  
 DB 1 TCAGCTTCGGGACTCGCTCTCTCTATCCAAACCGGAATAAATGGGACATCGTATG 60  
 QY 21 SerAlaMetIleProGluIleAsp 28  
 DB 61 TCGGCGATGATACACAGATCGAT 84

RESULT 3  
 ABK28625  
 ID ABK28625 standard; cDNA; 1506 BP.

XX ABK28625;

XX 09-APR-2002 (first entry)

XX cDNA encoding A. thaliana cytokinin oxidase AtCKX2.

XX Arabidopsis thaliana; plant; cytokinin oxidase; transgenic plant;  
 KW root growth; lateral root; adventitious root; root geotropism; herbicide;  
 KW root meristem; shoot meristem; leaf senescence; parthenocary; gene; ss.

XX Arabidopsis thaliana.

OS WO200196580-A2.

PN 20-DEC-2001.

XX 18-JUN-2001; 2001WO-EP006833.

XX 16-JUN-2000; 2000EP-00870132.

PR 27-DEC-2000; 2000US-0258415P.

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 18, 2005, 07:01:12 ; Search time 256 Seconds  
(without alignments)  
647.472 Million cell updates/sec

Title: US-10-014-101B-32

Perfect score: 147  
Sequence: 1 SASGLALLYPTNRKNDNRMSAMIPED 28

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Zgapop 6.0, Zgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2.1/USPTO.spool/US10014101/runat.16022005.075933.8289/app.query.fasta\_1.199  
-DB=N Geneseq 16Dec04 -QEMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blossum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USR=US10014101 @CGN 1 1 352 @runat.16022005.075933.8289 -NCPUS=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq\_16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	147	100.0	84	6	Abk28630 A. thaliana
2	147	100.0	84	9	Acc85299 Arabidops
3	147	100.0	1506	6	Abk28625 cDNA enco
4	147	100.0	1506	9	Acc85294 Arabidops
5	147	100.0	1506	10	Adh61267 Arabidops

6	147	100.0	1506	12	ADQ06498
7	130	88.4	1548	3	AAC43214
8	108	73.5	1575	3	AAC42983
9	108	73.5	1575	6	ABZ14284
10	108	73.5	1575	6	ABK28627
11	108	73.5	1575	8	ADA68648
12	108	73.5	1575	9	ACC85296
13	108	73.5	1575	10	ADH61269
14	108	73.5	1555	10	ADH61279
15	96.5	65.6	2291	6	ABK28607
16	96.5	65.6	2291	9	ACC85276
17	96	65.3	1572	6	ABK28626
18	96	65.3	1572	9	ACC85295
19	96	65.3	1572	10	ADH61268
20	88	59.9	1590	8	ADA69773
21	87	59.2	1577	8	ADA70177
22	87	59.2	2282	12	ADP03325
23	87	59.2	2302	12	ADP03322
24	85	57.8	1587	8	ADA69574
25	83	56.5	1566	8	ADA71115
26	83	56.5	2782	6	ABK28609
27	83	56.5	2782	9	ACC85278
28	80	54.4	1515	6	ABK28629
29	80	54.4	1515	9	ACC85298
30	79	53.7	482	13	ACN54881
31	79	53.7	484	13	ACN55634
32	79	53.7	788	13	ADR61296
33	77	52.4	1611	6	ABK28628
34	77	52.4	1611	9	ACC85297
35	77	52.4	1620	6	ABK28632
36	77	52.4	1620	9	ACC85301
37	77	52.4	1623	10	ADH61271
38	75	51.0	1857	10	ADH61281
39	69	46.9	1575	10	ADH61270
40	69	46.9	1873	10	ADH61280
41	68	46.3	413	10	ADH61283
42	68	46.3	413	10	ADH61282
C 43	68	46.3	573	13	ACN53055
C 44	68	46.3	581	13	ACN59325
45	63	42.9	1605	2	AAX02914

#### ALIGNMENTS

RESULT 1

ABK28630  
ID ABK28630 standard; cDNA; 84 BP.

XX  
AC ABK28630;

DT 09-APR-2002 (first entry)

XX A. thaliana cytokinin oxidase AtCKX2, cDNA fragment.

XX Arabidopsis thaliana; plant; cytokinin oxidase; transgenic plant;  
XX root growth; lateral root; adventitious root; root geotropism; herbicide;  
XX root meristem; shoot meristem; leaf senescence; parthenocary; gene; ss.

OS Arabidopsis thaliana.

XX WO200196580-A2.

XX 20-DEC-2001.

XX 18-JUN-2001; 2001WO-EP006833.

XX 16-JUN-2000; 2000EP-00870132.

PR 27-DEC-2000; 2000US-0258415P.

PR 16-MAR-2001; 2001EP-00870053.

PA (SCHM/) SCHMULLING T.

PA (WERN/) WERNER T.

XX

***This Page Blank (uspto)***

source 1. .342  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/variety="Columbia-0 NASC stock code N1092"  
/db\_xref="taxon:3702"  
/clone="AC005917"  
/notes="Derived from superpool 8.28 NASC code N40371"  
1. .342  
/standard\_name="SM\_3.25119"

STS  
ORIGIN

Alignment Scores:  
Pred. No.: 0.00233 Length: 342  
Score: 77.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 93.75% Mismatches: 0  
Query Match: 52.38% Indels: 0  
DB: 11 Gaps: 0

US-10-014-101B-32 (1-28) x BX255834 (1-342)

QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrp 16  
|||||  
DB 243 TCAGCTTCGGGACTCGCTCTCTCTATCCAAACCGGATAAGTAC 290  
|||||

RESULT 40  
BX255816  
LOCUS BX255816 366 bp DNA linear STS 11-JUN-2003  
DEFINITION Arabidopsis thaliana transposon insertion STS SM\_3.25117, sequence  
tagged site.  
ACCESSION BX255816  
VERSION BX255816.1 GI:28411106  
KEYWORDS STS; STS, sequence tagged site.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1  
AUTHORS Clarke, J.H., Bowles, B., Carter, J., Hart, D., McCullagh, B.,  
Murphy, G., Langham, S., LeGrys, C., Jones, J.D.G. and Bevan, M.  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 366)  
AUTHORS Clarke, J.H.  
TITLE Direct Submission  
JOURNAL Submitted (17-FEB-2003) Clarke J.H., John Innes Centre, Colney  
Lane, Norwich, NR4 7UJ, UK  
COMMENT AT denotes an activation tag dissociation transposon within a  
single line, Et an enhancer trap dissociation transposon, Gr a gene  
trap dissociation transposon, MT a mis-expression enhancer trap  
dissociation transposon, SM a defective suppressor mutator  
transposon. \_3 denotes a sequence derived from the 3' end of the  
transposon. \_5 denotes a sequence derived from the 5' end of the  
transposon. BSRG GARNet, ATIS project  
On-line seed stock requests: <http://nasc.nott.ac.uk/> NASC stock  
code: N111666.

FEATURES  
source  
1. .366  
Location/Qualifiers  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/variety="Columbia-0 NASC stock code N1092"  
/db\_xref="taxon:3702"  
/clone="AC005917"  
/notes="Derived from superpool 8.28 NASC code N40371"  
1. .366  
/standard\_name="SM\_3.25117"

STS  
ORIGIN

Alignment Scores:  
Pred. No.: 0.00252 Length: 366  
Score: 77.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 93.75% Mismatches: 0

Query Match: 52.38% Indels: 0  
DB: 11 Gaps: 0  
US-10-014-101B-32 (1-28) x BX255816 (1-366)  
QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrp 16  
|||||  
DB 243 TCAGCTTCGGGACTCGCTCTCTCTATCCAAACCGGATAAGTAC 290  
|||||

Search completed: February 18, 2005, 12:07:21  
Job time : 6634 secs

JOURNAL Submitted (06-APR-2001) Yang S., Department of Biological Sciences,  
National University of Singapore, 119260, SINGAPORE  
COMMENT On Apr 12, 2001 this sequence version replaced gi:11558276.  
FEATURES  
source

1. .1740  
/organism="Dendrobium cv. 'Sonia'"  
/mol\_type="mRNA"  
/varity="28"  
/db\_xref="taxon:136995"  
/notes="Dendrobium hybrid cultivar"  
1. .1740  
/genes="ckol"  
34. .1644  
/genes="ckol"  
/function="metabolic enzyme of cytokinins"  
/codon\_start=1  
/product="cytokinin oxidase"  
/protein\_id="CAC17752.1"  
/db\_xref="GI:11558277"  
/db\_xref="GOA:Q3PE45"  
/db\_xref="UniProt/TrEMBL:Q9FE45"  
/translation="MNLHAMPPLNPTSLTLLTTLMSILIQSPNSLPTNLLTHPTSTH  
LRFDSLSSAASDFDGIHSLPSAVFLPSPSDIATLLRSLHFSFHSFTVSGARGLGH  
STRGQAQFGIWINNPSIDGGITVSDIGMFDAGAEOMWIDVLRSLRHLTPKSWT  
DYLYLGGTSLNGGISGQAFHLGPOISNVHLDIVTGKMWVTCSESNPDLFSLV  
GGLGQGIITRARIALAKAPQSVRWMLYTDPELTKQOELLISIKARSGEWNLYV  
EGSLMHELSKNWRSPFFSEKDKIKKLAGNEGVITYCLASFYIDYGHENNSRA  
DRAQMDDEELRLKLSVSGFANDSVYMGFLNRVHGDGLKRAMGLWPHPLWN  
LFVSKNIDFHIGVFKGMKNSGPIIVPTKSKWDMKSTSIPEERVFYISGI  
LISSEMDLEHLESHNAELKFCDOQGMVQYLPHTYSIEDWKHFGKWERFVEMK  
SRYDPKAILSPGQKIFTHLVDELCLSDH"

gene

CDS

1. .1740  
/organism="Dendrobium cv. 'Sonia'"  
/mol\_type="mRNA"  
/varity="28"  
/db\_xref="taxon:136995"  
/notes="Dendrobium hybrid cultivar"  
1. .1740  
/genes="ckol"  
34. .1644  
/genes="ckol"  
/function="metabolic enzyme of cytokinins"  
/codon\_start=1  
/product="cytokinin oxidase"  
/protein\_id="CAC17752.1"  
/db\_xref="GI:11558277"  
/db\_xref="GOA:Q3PE45"  
/db\_xref="UniProt/TrEMBL:Q9FE45"  
/translation="MNLHAMPPLNPTSLTLLTTLMSILIQSPNSLPTNLLTHPTSTH  
LRFDSLSSAASDFDGIHSLPSAVFLPSPSDIATLLRSLHFSFHSFTVSGARGLGH  
STRGQAQFGIWINNPSIDGGITVSDIGMFDAGAEOMWIDVLRSLRHLTPKSWT  
DYLYLGGTSLNGGISGQAFHLGPOISNVHLDIVTGKMWVTCSESNPDLFSLV  
GGLGQGIITRARIALAKAPQSVRWMLYTDPELTKQOELLISIKARSGEWNLYV  
EGSLMHELSKNWRSPFFSEKDKIKKLAGNEGVITYCLASFYIDYGHENNSRA  
DRAQMDDEELRLKLSVSGFANDSVYMGFLNRVHGDGLKRAMGLWPHPLWN  
LFVSKNIDFHIGVFKGMKNSGPIIVPTKSKWDMKSTSIPEERVFYISGI  
LISSEMDLEHLESHNAELKFCDOQGMVQYLPHTYSIEDWKHFGKWERFVEMK  
SRYDPKAILSPGQKIFTHLVDELCLSDH"

gene

CDS

1. .1740  
/organism="Dendrobium cv. 'Sonia'"  
/mol\_type="mRNA"  
/varity="28"  
/db\_xref="taxon:136995"  
/notes="Dendrobium hybrid cultivar"  
1. .1740  
/genes="ckol"  
34. .1644  
/genes="ckol"  
/function="metabolic enzyme of cytokinins"  
/codon\_start=1  
/product="cytokinin oxidase"  
/protein\_id="CAC17752.1"  
/db\_xref="GI:11558277"  
/db\_xref="GOA:Q3PE45"  
/db\_xref="UniProt/TrEMBL:Q9FE45"  
/translation="MNLHAMPPLNPTSLTLLTTLMSILIQSPNSLPTNLLTHPTSTH  
LRFDSLSSAASDFDGIHSLPSAVFLPSPSDIATLLRSLHFSFHSFTVSGARGLGH  
STRGQAQFGIWINNPSIDGGITVSDIGMFDAGAEOMWIDVLRSLRHLTPKSWT  
DYLYLGGTSLNGGISGQAFHLGPOISNVHLDIVTGKMWVTCSESNPDLFSLV  
GGLGQGIITRARIALAKAPQSVRWMLYTDPELTKQOELLISIKARSGEWNLYV  
EGSLMHELSKNWRSPFFSEKDKIKKLAGNEGVITYCLASFYIDYGHENNSRA  
DRAQMDDEELRLKLSVSGFANDSVYMGFLNRVHGDGLKRAMGLWPHPLWN  
LFVSKNIDFHIGVFKGMKNSGPIIVPTKSKWDMKSTSIPEERVFYISGI  
LISSEMDLEHLESHNAELKFCDOQGMVQYLPHTYSIEDWKHFGKWERFVEMK  
SRYDPKAILSPGQKIFTHLVDELCLSDH"

ORIGIN

Alignment Scores:  
Pred. No.: 0.00481 Length: 1740  
Score: 80.00 Matches: 14  
Percent Similarity: 73.91% Conservative: 3  
Best Local Similarity: 60.87% Mismatches: 6  
Query Match: 54.42% Indels: 0  
DB: 8 Gaps: 0

US-10-014-101B-32 (1-28) x DSO294542 (1-1740)

QY 4 GlyLeuAlaLeuLeuTyProThrAanArgAsnLysTrpAspAanArgMetSerAlaMet 23  
Db 1288 GGTCCAACTACTGTTTATCTCTCAAAAAGATCAAAATGGGATAAAGGATGCTACTTCA 1347

QY 24 IleProGlu 26

Db 1348 ATTCACAT 1356

RESULT 38

AY209184

LOCUS

DEFINITION

AY209184

VERSION

KEYWORDS

SOURCE

ORGANISM

Hordeum vulgare

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Poaceae; Triticeae; Hordeum.

1 (bases 1 to 1866)

Galuszka, P. and Frebort, I.

Cytokinin dehydrogenase genes in barley

Unpublished

2 (bases 1 to 1866)

Galuszka, P. and Frebort, I.

Direct Submission

Submitted (02-JAN-2003) Dept. of Biochemistry, University of

Palacky, Slechtitelu 11, Olomouc 78371, Czech Republic

FEATURES

source

1. .1866  
/organism="Hordeum vulgare"  
/mol\_type="mRNA"  
/cultivar="Luxor"  
/db\_xref="taxon:4513"  
89. .1651  
/EC\_number="1.5.99.12"  
/codon\_start=1  
/product="cytokinin dehydrogenase 3"  
/protein\_id="AAO50082.1"  
/db\_xref="GI:28883584"  
/translation="MKQLLQVLLKLLGLSRVTTEHPVKYDVLASLCTLPDGHFS  
FHDLPAAARDFCNLSPFPVAVLHFGSVADIARTVRHFLMGEHSLTLTYAARGHSHL  
YQSQAGGIVIRMESLQSVQMHPGASPYVDSGGLWVNLVNLTKYGLAPKSWT  
DYHLTVGGTSLNAGVSGQTFRHGPOISNVNELEIVTGRDITCSPEQNSDLFHAAL  
GGLGQGIITRARIALAPQSVRWMLYTDPELTKQOELLISIKARSGEWNLYV  
INRTGILNNRSPFPQPARASQFETDKVLFLEMTKNFNEPAGIEQIHALLSQ  
LRYTPSLFHTDVTYMEFLDRVHSEIKLRAKGLMEVHPWMLNLIIPRTVHTFAKQV  
FGKILEDNNGPILLYPNKSRWDRTSVVLDPDEVSIVLGGFLPSAMGPHSIKRTLNL  
NNQILEFNKASIGVKQLPLHYSTBFEWKAHYGAWDAFQQRKNITDPLAILAQORI  
FQRTFASPLSS"

CDS

1. .1866  
/organism="Hordeum vulgare"  
/mol\_type="mRNA"  
/cultivar="Luxor"  
/db\_xref="taxon:4513"  
89. .1651  
/EC\_number="1.5.99.12"  
/codon\_start=1  
/product="cytokinin dehydrogenase 3"  
/protein\_id="AAO50082.1"  
/db\_xref="GI:28883584"  
/translation="MKQLLQVLLKLLGLSRVTTEHPVKYDVLASLCTLPDGHFS  
FHDLPAAARDFCNLSPFPVAVLHFGSVADIARTVRHFLMGEHSLTLTYAARGHSHL  
YQSQAGGIVIRMESLQSVQMHPGASPYVDSGGLWVNLVNLTKYGLAPKSWT  
DYHLTVGGTSLNAGVSGQTFRHGPOISNVNELEIVTGRDITCSPEQNSDLFHAAL  
GGLGQGIITRARIALAPQSVRWMLYTDPELTKQOELLISIKARSGEWNLYV  
INRTGILNNRSPFPQPARASQFETDKVLFLEMTKNFNEPAGIEQIHALLSQ  
LRYTPSLFHTDVTYMEFLDRVHSEIKLRAKGLMEVHPWMLNLIIPRTVHTFAKQV  
FGKILEDNNGPILLYPNKSRWDRTSVVLDPDEVSIVLGGFLPSAMGPHSIKRTLNL  
NNQILEFNKASIGVKQLPLHYSTBFEWKAHYGAWDAFQQRKNITDPLAILAQORI  
FQRTFASPLSS"

Alignment Scores:  
Pred. No.: 0.00766 Length: 1866  
Score: 79.00 Matches: 12  
Percent Similarity: 79.17% Conservative: 7  
Best Local Similarity: 50.00% Mismatches: 5  
Query Match: 53.74% Indels: 0  
DB: 8 Gaps: 0

ORIGIN

US-10-014-101B-32 (1-28) x AY209184 (1-1866)

QY 3 SerGlyLeuAlaLeuLeuTyProThrAanArgAsnLysTrpAspAanArgMetSerAla 22

Db 1292 AATGTCCTCATTTGCTCTACCCAGTGACAGTCCAGATGGACACCGACATCAGTG 1351

QY 23 MetIleProGlu 26

Db 1352 GTCTTACCAGAT 1363

RESULT 39

BX255834

LOCUS

DEFINITION

AX255834

VERSION

KEYWORDS

SOURCE

ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi

1

Clarke, J.H., Bowles, B., Carter, J.J., Hart, D., McCullagh, B.,

Murphy, G., Latham, S., LeGrys, C., Jones, J.D.G. and Bevan, M.

Unpublished

2 (bases 1 to 342)

Clarke, J.H.

Direct Submission

Submitted (17-FEB-2003) Clarke J.H., John Innes Centre, Colney

Lane, Norwich, NR4 7UJ, UK

AT denotes an activation tag dissociation transposon within a

single line, ET an enhancer trap dissociation transposon, GT a gene

trap dissociation transposon, MT a mis-expression enhancer trap

dissociation transposon, SM a defective suppressor mutator

transposon. 3 denotes a sequence derived from the 3' end of the

transposon, 5 denotes a sequence derived from the 5' end of the

transposon BSRG GARNET, ATIS project

On-line seed stock requests: <http://nasc.nott.ac.uk/> NASC stock

code: N112982

Location/Qualifiers

FEATURES

Qy	21	SerAlaMetIleProGlu	26
		:::	
Db	1210	TCAGCAGTAACCCGGAG	1227
 RESULT 36			
AX785081	AX785081	Sequence 30 from Patent WO03050287.	linear PAT 17-JUL-2003
LOCUS	AX785081	1515 bp	DNA
DEFINITION	AX785081		
ACCSSION	AX785081		
VERSION	AX785081.1	GI:32952912	
KEYWORDS			
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		

```

REFERENCE
AUTHORS      Schumullig, T. and Werner, T.
TITLE        Method for modifying plant morphology, biochemistry and physiology
JOURNAL      Patent: WO 03050287-A 30 19-JUN-2003;
              Schumullig, Thomas (DE) ; Werner, Tomas (DE)
FEATURES
source       1..1515
              /organism="Arabidopsis thaliana"
              /mol_type="unassigned DNA"
              /db_xref="taxon:3702"

ORIGIN

Alignment Scores:
Pred. No.:    0.0041
Score:        80.00
Percent Similarity: 80.77%
Best Local Similarity: 46.15%
Query Match:  54.42%
              Length: 1515
              Matches: 12
              Conservative: 9
              Mismatches: 5
              Indels: 0

```

DB:		6	Gaps:	0
US-10-014-101B-32 (1-28) x AY785081 (1-1515)				
Qy	1	SerAlaSerGlyLeuAlaLeuLeuTyPrThrAsnArgAsnLysTrpIePAsnArgMet 20		
		:::		
Db	1150	ACAAGCAACGGCCGCTCGTCTACCCAGTGAACAAATCAAAGTGGGCAATCAAACA 1209		
Qy	21	SerAlaMetIleProGlu 26		
		:::		
Db	1210	TCACGAGTACACCGGAG 1227		
RESULT 37				
DSO294542				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
DSO294542				
Dendrobium 'Sonia' mRNA linear PLN 10-APR-2001				
Dendrobium 'Sonia' mRNA for cytokinin oxidase (ckol gene).				
AT294542				
AJ294542.2 GI:13620164				
ckol gene; cytokinin oxidase.				
Dendrobium cv. 'Sonia'				
Dendrobium cv. 'Sonia'				
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
Spermatophyta; Magnoliopsida; Liliopsida; Asparagales; Orchidaceae;				

REFERENCE AUTHORS	Yang, S., Yu, H. and Goh, C.
TITLE	Molecular cloning and characterization of a cDNA encoding cytokinin oxidase in <i>Dendrobium sonia</i> orchid
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 1740)
AUTHORS	Yang, S.
TITLE	Direct Submission
JOURNAL	Submitted (11-SEP-2000) Yang S., Department of Biological Sciences, National University of Singapore, 119260, SINGAPORE
REMARK	revised by [3]
REFERENCE	3 (bases 1 to 1740)
AUTHORS	Yang, S.



AUTHORS	Rose,M., Hempel,S., Entian,K.-D., Mewes,H.W., Lemcke,K. and Mayer,K.F.X.
JOURNAL	Unpublished
REFERENCE	3 (bases 110074 to 200000)
AUTHORS	Zimmermann,W., Gruenewald,A., Wambutt,R., Kalicki,J., Wohldmann,P., Smith,A., Mewes,H.W., Lemcke,K. and Mayer,K.F.X.
JOURNAL	Unpublished
REFERENCE	4 (bases 1 to 200001)
AUTHORS	EU Arabidopsis sequencing, project.
TITLE	Direct Submission
JOURNAL	Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: lemcke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bbsrc.ac.uk
COMMENT	Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: <a href="http://www.mips.biochem.mpg.de/proj/thal/">http://www.mips.biochem.mpg.de/proj/thal/</a> this fragment has an overlap with ATCHRIV70 at the 5' end and an overlap with ATCHRIV72 at the 3' end.
FEATURES	<p>Location/Qualifiers</p> <p>1..200001</p> <p>/organism="Arabidopsis thaliana"</p> <p>/mol_type="genomic DNA"</p> <p>/variety="Columbia"</p> <p>/db_xref="taxon:3702"</p> <p>/chromosome="4"</p> <p>5366..6855</p> <p>/gene="AT4g29420"</p> <p>complement(join(5366..6258,6408..6855))</p> <p>/gene="AT4g29420"</p> <p>complement(join(5366..6258,6408..6855))</p> <p>/gene="AT4g29420"</p> <p>/note="contains EST gb:F19848"</p> <p>/codon_start=1</p> <p>/product="hypothetical protein"</p> <p>/protein_id="CAB79700.1"</p> <p>/db_xref="GI:7269841"</p> <p>/db_xref="UniProt/TREMBL:Q9M0E1"</p> <p>/translation="MDPELPKILSRINDESLARCVASRTKLSISREVRANL ICTWSYLRKSRIVVTPFKTIERSLIENSSKIRISVGVKALKGMSFDDNEEDSK DLYLTVDFVEMKPLRYVRELDNISDFWQCRKSDILALISNCKSLVKLVKN AWLVSVDVFWMLRYLLEFIRLDNLEKNCDFPQELNLIGVGLGRIHFL HLKSCWTVSNAPLSIAIYAPNLELKLCKNPKSLLVETPLVQCHLSVEDAGVSF GFQDQKLTLELVSDPMVRLISNIFGNKIKLAVDSVKSTEQSERLELGLATLKAPP GITSLSLSPRTWSDIRTFQSCGLGDMKGTDSLKQITARVQMSDHTNVHQTVSFIRS IVNKYRGLEDMRLMIHQDKDPRVRSNLSICTMNSRVRVKWGWAEGGEDMVV"</p> <p>complement(5366..6258)</p> <p>/gene="AT4g29420"</p> <p>/number=1</p> <p>complement(6259..6407)</p> <p>/gene="AT4g29420"</p> <p>/number=1</p> <p>6299..6328</p> <p>complement(6408..6855)</p> <p>/gene="AT4g29420"</p> <p>/number=2</p> <p>7091..8034</p> <p>/gene="AT4g29430"</p> <p>join(7091..7220,7698..7858,7936..8034)</p> <p>/gene="AT4g29430"</p> <p>/note="EST GB:R83969 marks first intron</p> <p>strong similarity to ribosomal protein S15a - Brassica napus, PIR2:520945</p> <p>Contains Ribosomal protein S8 signature AA99-116</p> <p>contains EST gb:AW004167.1, AA395604, R83969"</p> <p>/codon_start=1</p> <p>/product="ribosomal protein S15a homolog"</p> <p>/protein_id="CAB79701.1"</p> <p>/db_xref="GI:7269842"</p> <p>/db_xref="GOA:Q9M0E0"</p> <p>/db_xref="UniProt/TREMBL:Q9M0E0"</p>
gene	
gene	
CDS	
exon	<p>/translation="MGRRILNDRALRTVNAERRGKASVELKPISTVMSSFLRMKEKG YIKNFQVYDHPHVRIGRTIVDQGRVNDCKALTYQDVRAKEIEKYTERLTLPTRQWGYV ITTPDGIILDHEEAIKRNVGQVLGFFY"</p> <p>7091..7220</p> <p>/gene="AT4g29430"</p> <p>/number=1</p> <p>7221..7697</p> <p>/gene="AT4g29430"</p> <p>/number=1</p> <p>7698..7858</p> <p>/gene="AT4g29430"</p> <p>/number=2</p> <p>7859..7935</p> <p>/gene="AT4g29430"</p> <p>/number=2</p> <p>7936..8034</p> <p>/gene="AT4g29430"</p> <p>/number=3</p> <p>8593..12224</p> <p>/gene="AT4g29440"</p> <p>complement(join(8593..11057,11144..11363,11456..11593,11689..11860,11933..12036,12108..12224))</p> <p>/gene="AT4g29440"</p> <p>complement(join(8593..11057,11144..11363,11456..11593,11689..11860,11933..12036,12108..12224))</p> <p>/gene="AT4g29440"</p> <p>/note="similarity to putative suppressor protein - Arabidopsis thaliana, PID:g3687246 contains EST gb:Z34655"</p> <p>/codon_start=1</p> <p>/product="putative protein"</p> <p>/protein_id="CAB79702.1"</p> <p>/db_xref="GI:7269843"</p> <p>/db_xref="GOA:Q9M0D9"</p> <p>/db_xref="UniProt/TREMBL:Q9M0D9"</p> <p>/translation="MAASLKILKNKQTOIKQLRRELAHLLESQOTOTAKIRVHVH REEKTVAAELVGIYCELLVLRGLVIDSQKTPNDLKEAVSVLYASQLRTDVGELSD IVKHFSAKYKDFVSAAGLQDPDSGVSRLLVEKLSVKAPDGPTKIKILTEIATQHNT WEAESLIVSDPKETMSAGASSSVQATPKIKESRIQNNQPPVQAAATVNVQNS VATDGRSSRTSTDFNVGKTPDHVQDPKSGDRVDGRHDSHNSYENLRNPHS VATAAARAAESAEARAAELSKNNRQVDQSTTTTAAESSKKTVDLSENTSWRHGHSREN RTSSNMGGKGGKELLKNNRQVDQSTTTTAAESSKKTVDLSENTSWRHGHSREN SLEMRNPSFAGIKREKQKQDDNLSSADVLNKKQSSRASHSPSSNFSDDNDVT ALHDIDSPSIFENRQSTVDRESYNDSPVVVAFADYSSFPDKPQDTEADAHD EPEQGLSPALSGSSKTSKTDHMPTEISSWSLEGLKDLGLSSASTSQVLKKEKPSPT PDLCRYPPSDTEQDDSKTOBERSDAETPTGLKFGPLASGLNLTLPVSGSSPPR DTDGTSPPASLHPEPSAKFDYDRSEEDNLRLSGRAEGSKLTAQKSHMSEG DTSSKSIKEYLPTVDPSRSSLSLATASSSSIRNELYTKASNDKRPSSIIPDSSSPR DESDNELPKRVFYQEKRTSRTRPHTLHSGVSHKLEELIPTRASTRSQDRTHK TTPASASAFYHTSSDDDEKEVHRDTHAIQTRPYTISIRRTKGQERRPSLVTKLID KVSFDESPPLKSPKALPTKQGSASSLSYLPKTEKVKSHDQESHPLGLGAKPLIKQ QGSASSLSPLTNKASPDQSPPLVPEKPEAKQKQSGASSLSFLPKTDKASPDQDS PPKLLPKEKPAKQGSATSSSSLPKTEKISHYRSPSKLTPPEAKSMAKQEGGLASSSS SLPKTVTSPDPTPAKEKASHVHKLDPDYDDIFAKLGLRR"</p> <p>complement(8593..11057)</p> <p>/gene="AT4g29440"</p> <p>/number=1</p> <p>complement(11058..11143)</p> <p>/gene="AT4g29440"</p> <p>/number=1</p> <p>complement(11144..11363)</p> <p>/gene="AT4g29440"</p> <p>/number=2</p> <p>complement(11364..11455)</p> <p>/gene="AT4g29440"</p> <p>/number=2</p> <p>complement(11456..11593)</p> <p>/gene="AT4g29440"</p> <p>/number=3</p> <p>complement(11594..11688)</p> <p>/gene="AT4g29440"</p> <p>/number=3</p> <p>complement(11689..11860)</p>
intron	
exon	
intron	
exon	
intron	
exon	
intron	
exon	
intron	
exon	

```
/db_xref="GI:5123545"
/db_xref="UniProt/TREMBL:Q9SU94"
/translation="MTKNGNHDEEPISEPNQNTKIRFEDADEDEVARSGVAGBET
PQDESFDGASADTAETVDDTTSDYFDSYSHFGIHEMLKDVVVRTKTYQNVYQY
KFLIKDKILVDGAGTILSLFCAKAGAAHVAVGCSOMADMAKEIVKANGSDVITV
LKGKIEELPFPKVDVIISEWNGYELLFENMLDSVLVARDKWLVEGGVVLDPKASLH
LTAISEYKEDIKIEFVNSVYGFDMSCIKKKAMMELPVDVQONQIVTDSRLKTMDI
SRMSGDASTAPFLVAFQNRNDYIHALVAFVDFVFTMCHKLGLFGFSGPKSRATHWKQT
VLYLEDVLTICGETITGTMSVSPNKNPRDIDIKLSYSLNGQHCXISRTQHYQNR"
1575..1812
/gene="Tl6L4.20"
/number=1
1813..1907
/gene="Tl6L4.20"
/number=1
1908..1918
/gene="Tl6L4.20"
/number=2
1919..2040
/gene="Tl6L4.20"
/number=2
2041..2196
/gene="Tl6L4.20"
/number=3
2197..2292
/gene="Tl6L4.20"
/number=3
2293..2356
/gene="Tl6L4.20"
/number=4
2357..2436
/gene="Tl6L4.20"
/number=4
2437..2579
/gene="Tl6L4.20"
/number=5
2580..2669
/gene="Tl6L4.20"
/number=5
2670..2757
/gene="Tl6L4.20"
/number=6
2758..2864
/gene="Tl6L4.20"
/number=6
2865..2980
/gene="Tl6L4.20"
/number=7
2981..3070
/gene="Tl6L4.20"
/number=7
3071..3221
/gene="Tl6L4.20"
/number=8
3222..3382
/gene="Tl6L4.20"
/number=8
3383..3588
/gene="Tl6L4.20"
/number=9
3589..3493
/gene="Tl6L4.30"
complement(3892..4080,4202..4333,4424..4501,
4614..4745,4855..4923,5083..5193,5284..5493))
/gene="Tl6L4.30"
complement(join(3892..4080,4202..4333,4424..4501,
4614..4745,4855..4923,5083..5193,5284..5493))
/notes="contains EST gb:H76475"
/codon_start=1
/product="putative protein"
/protein_id="CAB45312.1"
```

```
/db_xref="GI:5123546"
/db_xref="UniProt/TREMBL:Q9SU93"
/translation="MVVNRVLVALTLLASSALLPVSDAAKPSSTPRKEDVPYIK
QOVCKLSRLHLVKRQQQKISPKKI SEYEIIEIAENVNLKKEADWMLKIDIVEK
GDNVLVQOQEGMCSKKTENACQKVI GYSIDTVABYIYKSPDLVSLVNLCKD
LTDACSKPPVPKORVPEPFAKPSKDAEMDKILRSQMGMPGAPGMKVYSREIDIEK
GNIGNDDDDGDEDEEDDKFPKNLGVLEKESKTEBELKKTITKEPKKGEALKRH
AQVSNRVRVWVGSSSSKPKSKGSEL"
complement(3892..4080)
/gene="Tl6L4.30"
/number=1
complement(4081..4201)
/gene="Tl6L4.30"
/number=1
complement(4202..4333)
/gene="Tl6L4.30"
/number=2
complement(4334..4423)
/gene="Tl6L4.30"
/number=2
complement(4424..4501)
/gene="Tl6L4.30"
/number=3
complement(4502..4613)
/gene="Tl6L4.30"
/number=3
complement(4614..4745)
/gene="Tl6L4.30"
/number=4
complement(4746..4854)
/gene="Tl6L4.30"
/number=4
complement(4855..4923)
/number=4

Alignment Scores:
Pred. No.: 0.141 Length: 88011
Score: 83.00 Matches: 19
Percent Similarity: 37.93% Conservative: 3
Best Local Similarity: 32.76% Mismatches: 6
Query Match: 56.46% Indels: 30
DB: 8 Gaps: 1

US-10-014-101b-32 (1-28) x ATT16L4 (1-88011)

QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLys----- 15
Db 78543 TCAACTTCTGTGTACTCTCTCTCTATCCACAAACCGAAAGTAATAATTACTTTT 78602
QY 15 ----- 15

Db 78603 TGATTTTGTATTATTGAAGTATATCCAAATAATGTATGTTAAATTGTTAACAGAAAT 78662
QY 16 -----TipAspAsnArgMetSerAlaMetIleProGluIleAsp 28
Db 78663 TATTTTATTATAGATGGACACCGCATGTCAACGATGACACCGACGACGAGAT 78716

RESULT 33
ATCHRIV71 ATCHRIV71 200001 bp DNA linear PLN 16-MAR-2000
LOCUS Arabidopsis thaliana DNA chromosome 4, contig fragment No. 71.
DEFINITION Arabidopsis thaliana (thale cress)
ACCESSION AL161575
VERSION AL161575.2 GI:7269840
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 24920)
Lennard,N., Quail,M., Harris,B., Rajandream,M.A., Barrell,B.G.,
Mewes,H.W., Lemcke,K. and Mayer,K.F.X.
JOURNAL
Unpublished
REFERENCE
2 (bases 24816 to 112826)
```

```

Qy 16 -----TTPAspAsnArgMetSerAlaMetIleProGluIleAsp 28
Db 2477 TATTTTATTATAGTGAACAACCGCATGTCAACGATGACACCGGACGAAGAT 2530

RESULT 31
LOCUS AX785058 2782 bp DNA linear PAT 17-JUL-2003
DEFINITION Sequence 7 from Patent WO03050287.
ACCESSION AX785058
VERSION AX785058.1 GI:32952892
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE 1
AUTHORS Schmulling, T. and Werner, T.
TITLE Method for modifying plant morphology, biochemistry and physiology
JOURNAL Patent: WO 03050287-A 7 19-JUN-2003;
Schmulling, Thomas (DE); Werner, Tomas (DE)
FEATURES
source
1. .2782
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Alignment Scores:
Pred. No.: 0.00261 Length: 2782
Score: 83.00 Matches: 19
Percent Similarity: 37.93% Conservative: 3
Best Local Similarity: 32.76% Mismatches: 6
Query Match: 56.46% Indels: 30
DB: 6 Gaps: 1

US-10-014-101B-32 (1-28) x AX785058 (1-2782)

Qy 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLys----- 15
Db 2357 TCAACTTCGTGTGTACTCTCTCTATCCCAACGAAACGAAGTAAATATTACTTTT 2416
Qy 15 ----- 15
Db 2417 TGATTTGTTTATTGAAAGTATATCCCAATATGATGTTAAATGTTTAAACAAGATT 2476
Qy 16 -----TTPAspAsnArgMetSerAlaMetIleProGluIleAsp 28
Db 2477 TATTTTATTATAGTGAACAACCGCATGTCAACGATGACACCGGACGAAGAT 2530

RESULT 32
LOCUS ATT16L4 88011 bp DNA linear PLN 17-JUN-1999
DEFINITION Arabidopsis thaliana DNA chromosome 4, BAC clone T16L4 (BSSA
project).
ACCESSION AL079344
VERSION AL079344.1 GI:5123543
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE 1
AUTHORS Bevan, M., Rose, M., Hempel, S., Entian, K.-D., Bancroft, I.,
Mewes, H.W., Mayer, K.F.X., Lemcke, K. and Schueller, C.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 88011)
AUTHORS EU Arabidopsis sequencing, project.
TITLE Direct Submission
Submitted (17-JUN-1999) MIPS, at the Max-Planck-Institut fuer
Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
schueller@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de Project

```

Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,  
E-mail: michael.bevan@bbc.ac.uk  
Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/thal/>.

Location/Qualifiers  
1. .88011  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/variety="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="4"  
47. .1157  
/gene="T16L4.10"  
join(<47..117,312..365,446..551,648..716,803..946,  
1056..1157)  
/gene="T16L4.10"  
/note="similarity to X-Pro dipeptidase, Homo sapiens,  
PIR2:A32454"  
/codon\_start=1  
/product="X-Pro dipeptidase-like protein (fragment)"  
/protein\_id="CAB45310.1"  
/db\_xref="GI:5123544"  
/db\_xref="GOA:Q9SU95"  
/db\_xref="UniProt/TREMBL:Q9SU95"  
/translation="AVLDARNSVLSAMKPGVNWDMHKLAEKILLESILKSGILTGDV  
DDMVQRLGAVPMHGLGHFMGIDHTDGGYKGVVERPKFKGLKSLRTARDLGGNVI  
TVPCGYFIKALLFPAMANATTSKFFNRTIERFNFVGVRIESDLVVVTANGCKNMNTN  
VPRETWEIEAVWAGGPPPTK"  
47. .117  
/gene="T16L4.10"  
/number=1  
118. .311  
/gene="T16L4.10"  
/number=1  
312. .365  
/gene="T16L4.10"  
/number=2  
366. .445  
/gene="T16L4.10"  
/number=2  
446. .551  
/gene="T16L4.10"  
/number=3  
552. .647  
/gene="T16L4.10"  
/number=3  
648. .716  
/gene="T16L4.10"  
/number=4  
717. .802  
/gene="T16L4.10"  
/number=4  
803. .946  
/gene="T16L4.10"  
/number=5  
947. .1055  
/gene="T16L4.10"  
/number=5  
1056. .1157  
/gene="T16L4.10"  
/number=6  
1157. .3588  
/gene="T16L4.20"  
join(1575..1812,1908..1918,2041..2196,2293..2356,  
2437..2579,2670..2757,2865..2980,3071..3221,3383..3588)  
/gene="T16L4.20"  
/note="contains EST gb:H37349, T45910, H37681, Z26403,  
AA041059"  
/codon\_start=1  
/product="arginine methyltransferase (paml)"  
/protein\_id="CAB45311.1"





Db 1282 TCAGTGGTCATACAGAT 1299

RESULT 24  
AK121317  
LOCUS  
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J023113D01, full insert sequence.

ACCESSION AK121317  
VERSION AK121317.1 GI:37990940  
KEYWORDS FLI cDNA; CAP trapper.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE 1  
AUTHORS The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team; Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Ootomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,N., Xie,Q., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Miura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN, Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M., and Hayashizaki,Y.  
COLLECTION, mapping, and annotation of over 28,000 cDNA clones from japonica rice

TITLE Science 301 (5631), 376-379 (2003)  
JOURNAL Science 301 (5631), 376-379 (2003)  
MEDLINE 22752273  
PUBMED 12869764

REFERENCE 2  
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,P., Hotta,I., Iida,J., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakahama,Y., Nakamura,M., Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osato,N., Ota,Y., Ootomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tanaka,T., Tomaru,A., Toyota,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S., and Yoshimura,A.  
Collection, mapping, and annotation of 28K full-length cDNA clones from japonica rice

TITLE Unpublished  
JOURNAL Unpublished  
REFERENCE 3 (bases 1 to 2328)  
AUTHORS Kikuchi,S.  
TITLE Direct Submission  
JOURNAL Submitted (31-JAN-2003) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skkuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)  
COMMENT This clone is one of the 32K full-length cDNA clones from japonica rice.

URL : http://cdna01.dna.affrc.go.jp/cDNA/NIAS Rice Full-length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Yamamoto,M., and Nakahama,Y.  
FAIS Genome Sequencing & Analysis Group: Ootomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K., and Murakami,K.  
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraka,T., Hori,P., Iida,J., Imamura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toyota,T., Waki,K., Yasunishi,A., and Hayashizaki,Y.  
Location/Qualifiers

FEATURES  
source 1..2328  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="J023113D01"

ORIGIN  
Alignment Scores:  
Pred. No.: 0.000311 Length: 2328  
Score: 88.00 Matches: 16  
Percent Similarity: 73.08% Conservative: 3  
Best Local Similarity: 61.54% Mismatches: 7  
Query Match: 59.86% Indels: 0  
DB: 8 Gaps: 0

US-10-014-101b-32 (1-28) x AK121317 (1-2328)

QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnIysTrpAspMet 20  
Db 1449 AGCAACAATGGTCCCATCTGCTTTTACCAGTGAACAGAACCAAGTGGGACAGACA 1508  
QY 21 SerAlaMetIleProGlu 26  
Db 1509 TCAGTGGTCATACAGAT 1526

RESULT 25  
AX653630  
LOCUS AX653630  
DEFINITION Sequence 3500 from Patent WO03000898.  
ACCESSION AX653630  
VERSION AX653630.1 GI:29156444  
KEYWORDS Oryza sativa  
SOURCE Oryza sativa  
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE 1  
AUTHORS Chang,H.S., Chen,W., Cooper,B., Glazebrook,J., Goff,S.A., Hou,Y.M., Katagiri,F., Quan,S., Tao,Y., Whitham,S., Xie,Z., Zhu,T. and Zou,G.  
TITLE Plant genes involved in defense against pathogens  
JOURNAL Patent: WO 03000898-A 3500 03-JAN-2003;  
SYNOPSIS Syngenta Participations AG (CH)  
FEATURES Location/Qualifiers



```
mRNA      complement(join<28356..28814,28901..28989,29085..29148,
29244..29326,29422..29548,29658..29742,29822..29885,
29979..30040,30154..30224,30310..30405,30479..30627,
30708..30777,30847..30887,30959..31273))
/gene="At2g19470"
CDS       complement(join(28590..28814,28901..28989,29085..29148,
29244..29326,29422..29548,29658..29742,29822..29885,
29979..30040,30154..30224,30310..30405,30479..30627,
30708..30777,30847..30887,30959..31273))

Alignment Scores:
Pred. No.:      0.000832      Length:      92822
Score:          96.50         Matches:    27
Percent Similarity: 48.21%    Conservative: 0
Best Local Similarity: 48.21% Mismatches:    1
Query Match:     65.65%      Indels:     29
DB:              8           Gaps:      1

US-10-014-101B-32 (1-28) x AC005917 (1-92822)

QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLys----- 15
Db 39468 TCAGCTTCGGGACTCGCTCTCTCTATCAACAAACCGGAATAA-GTACATACCTCTCTT 39410
QY 15 ----- 15
Db 39409 CATTCAATTTATCTTCAAGAACCAAGTAATAATTTCTATGAAGTGAATGATGCTGTT 39350
QY 16 -----TrpAspAsnArgMetSerAlaMetIleProGluIleAsp 28
Db 39349 ATTGTTAGATGGGACAACTCGATGTCGGCGATGATACAGAGATCGAT 39302

RESULT 19
AX339730
LOCUS      AX339730          1572 bp      DNA      linear      PAT 10-JAN-2002
DEFINITION Sequence 27 from Patent WO0196580.
ACCESSION  AX339730
VERSION     AX339730.1  GI:18135723
KEYWORDS
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM    Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
AUTHORS     Schmulling,T. and Werner,T.
TITLE       Method for modifying plant morphology, biochemistry and physiology
JOURNAL     Patent: WO 0196580-A 27 20-DEC-2001;
Schmulling, Thomas (DE) ; Werner, Tomas (DE)
FEATURES    Location/Qualifiers
             source
               1..1572
               /organism="Arabidopsis thaliana"
               /mol_type="unassigned DNA"
               /db_xref="taxon:3702"

ORIGIN
Alignment Scores:
Pred. No.:      9.11e-06      Length:      1572
Score:          96.00         Matches:    18
Percent Similarity: 80.77%    Conservative: 3
Best Local Similarity: 69.23% Mismatches:    5
Query Match:     65.31%      Indels:     0
DB:              6           Gaps:      0

US-10-014-101B-32 (1-28) x AX339730 (1-1572)

QY 3 SerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22
Db 1231 ACCGGTCTGTTCTGTTTATCTATGATCGCAACAGTGAATGATCGATGCTGCC 1290

RESULT 20
AX339730
LOCUS      AX339730          1572 bp      mRNA      linear      PLN 08-NOV-2000
DEFINITION Arabidopsis thaliana cytokinin oxidase (CKX3) mRNA, complete cds.
ACCESSION  AF303979
VERSION     AF303979.1  GI:11120509
KEYWORDS
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM    Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
AUTHORS     Bilyeu,K.D., Laskey,J.G., Riekhof,W.R., VanVickle,S. and
Morris,R.O.
TITLE       A family of cytokinin oxidases from Arabidopsis thaliana
JOURNAL     Unpublished
REFERENCE   2
AUTHORS     Bilyeu,K.D., Laskey,J.G., Riekhof,W.R., VanVickle,S. and
Morris,R.O.
TITLE       Direct Submission
JOURNAL     Submitted (11-SEP-2000) University of Missouri, 210 Waters Hall,
Columbia, MO 65211, USA
FEATURES    Location/Qualifiers
             source
               1..1572
               /organism="Arabidopsis thaliana"
               /mol_type="mRNA"
               /db_xref="taxon:3702"
               /chromosome="V"
             gene
               1..1572
               /gene="CKX3"
             CDS
               1..1572
```

```
AX785078
LOCUS      AX785078          1572 bp      DNA      linear      PAT 17-JUL-2003
DEFINITION Sequence 27 from Patent WO03050287.
ACCESSION  AX785078
VERSION     AX785078.1  GI:32952909
KEYWORDS
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM    Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
AUTHORS     Schmulling,T. and Werner,T.
TITLE       Method for modifying plant morphology, biochemistry and physiology
JOURNAL     Patent: WO 03050287-A 27 19-JUN-2003;
Schmulling, Thomas (DE) ; Werner, Tomas (DE)
FEATURES    Location/Qualifiers
             source
               1..1572
               /organism="Arabidopsis thaliana"
               /mol_type="unassigned DNA"
               /db_xref="taxon:3702"

ORIGIN
Alignment Scores:
Pred. No.:      9.11e-06      Length:      1572
Score:          96.00         Matches:    18
Percent Similarity: 80.77%    Conservative: 3
Best Local Similarity: 69.23% Mismatches:    5
Query Match:     65.31%      Indels:     0
DB:              6           Gaps:      0

US-10-014-101B-32 (1-28) x AX785078 (1-1572)

QY 3 SerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22
Db 1231 ACCGGTCTGTTCTGTTTATCTATGATCGCAACAGTGAATGATCGATGCTGCC 1290

QY 23 MetIleProGluIleAsp 28
Db 1291 GCTATACCGGAGAGAT 1308

RESULT 21
AF303979
LOCUS      AF303979          1572 bp      mRNA      linear      PLN 08-NOV-2000
DEFINITION Arabidopsis thaliana cytokinin oxidase (CKX3) mRNA, complete cds.
ACCESSION  AF303979
VERSION     AF303979.1  GI:11120509
KEYWORDS
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM    Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
AUTHORS     Bilyeu,K.D., Laskey,J.G., Riekhof,W.R., VanVickle,S. and
Morris,R.O.
TITLE       A family of cytokinin oxidases from Arabidopsis thaliana
JOURNAL     Unpublished
REFERENCE   2
AUTHORS     Bilyeu,K.D., Laskey,J.G., Riekhof,W.R., VanVickle,S. and
Morris,R.O.
TITLE       Direct Submission
JOURNAL     Submitted (11-SEP-2000) University of Missouri, 210 Waters Hall,
Columbia, MO 65211, USA
FEATURES    Location/Qualifiers
             source
               1..1572
               /organism="Arabidopsis thaliana"
               /mol_type="mRNA"
               /db_xref="taxon:3702"
               /chromosome="V"
             gene
               1..1572
               /gene="CKX3"
             CDS
               1..1572
```

```

/gene="At2g19420"
/notes="synonym: F3p11.2; predicted by genscan"
/join(<6440..6458,6961..7071,8318..>8622)
/gene="At2g19420"
/join(6440..6458,6961..7071,8318..8622)
/gene="At2g19420"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAD10141.1"
/db_xref="GI:4191772"
/translation="MHEKPKTKTYDELXNGYLAIYNGYCDLRYFTVKLENDICVSLV
WNTGMYTKSGKRIRINRYRPIAKAKKKPKTKEQLMDEPFSDEDVLTSLGFDG
GYIIPFKSLMLINIPDSFRDELQTMHVSFYLRREVGLKMPG"
complement(7164..7241)
/rpt_family="(TA)n"
complement(9348..9369)
/rpt_family="AT_rich"
9624..9757
/rpt_family="AT_rich"
9956..12479
/gene="At2g19430"
/notes="synonym: F3p11.3; contains a Tip Asp (WD 40) repeat
signature"
/join(<9956..10281,10369..10445,10527..10572,10700..>10760,
10849..10902,10993..11073,11165..11204,11477..>11577,
11750..11860,12390..>12479)
/gene="At2g19430"
/join(9956..10281,10369..10445,10527..10572,10700..>10760,
10849..10902,10993..11073,11165..11204,11477..>11577,
11750..11860,12390..>12479)
/gene="At2g19430"
/codon_start=1
/product="putative WD-40 repeat protein"
/protein_id="AAD10142.1"
/db_xref="GI:4191773"
/translation="MTGDAINWDEYRESILKERIETRTVFTAWAPARISNPDA
FVVASDGLTAFHSLNSVQSAGFYSGQDVMVAEPERVVRAHEGPAIDVKFYGED
EDALLSCDGDGVRGKVRFAESDVSLHLENHLKPLLELINFQHKGPWALSMP
EINAMVPSQSSVFTAAAGSCAYCWDVESGKIWTFKGSHDYLHTVSRSSASQILT
GSDGTARTWDCKTGKCVKVGISQDKKSLRVSSMALDGSBMLVCGQGNLALNWL
ASECVOTIPPAHVQDVMPEQKQVAVGCGYGVIVDSIQGSHLCTFRSSSL"
complement(12903..14545)
/gene="At2g19440"
/notes="synonym: F3p11.4"
complement(join(<12903..13516,13584..>14083,14223..>14545))
/gene="At2g19440"
complement(join(12903..13516,13584..>14083,14223..>14545))
/gene="At2g19440"
/codon_start=1
/product="putative beta-1,3-glucanase"
/protein_id="AAD10143.1"
/db_xref="GI:4191774"
/translation="WNLLAFVVGIMGVIMVDGLGVNKGTMATHKLPPKVVQMLKD
NNINKYLFDAETTHSALSGLGVNVAIPNDQLKVMGSYDRAKDWHKVNTRNFN
GGVNIITFVAGPEPFLKSYNGSFINTFPALQINQALNEAGLSVSKATVPLNADVY
DPSNPFSAGRPDIIGQMTIVDFLGNNSAPITINITYPLSLYGNDDPLNYAF
FQCAKVDNAGIAYTFDANFTLVLSALKAVGHDMPITIVGVEGWPTEGDKHANS
AVRFYNGLLPRLGENRGTPRLPTYIEVYLFGLLDDAKSIAPCEPERHWGIFKFGQSP
KEPIDLSGQGNKLLIASENTVYQPKKCMNTEAKDLTKLAANIDYACTFSDCTALG
YSSCNTLQANGASAFNNYQVKNQDDEACIFQGLATITTKNISQGCNFPQIQA
STASSPSSSLVLLIAGVWFLLSGVMEV"
15295..17666
/rpt_family="Rf:ATREPS|ATREPS An AT-rich repetitive
sequence - a consensus."
15590..15728
/rpt_family="AT_rich"
15912..16149
/rpt_family="AT_rich"
complement(16223..16264)
/rpt_family="AT_rich"
16486..16577
/rpt_family="AT_rich"
complement(16668..16701)

```

```

/rpt_family="AT_rich"
16808..16936
/rpt_family="AT_rich"
complement(17023..17050)
/rpt_family="AT_rich"
17139..17462
/rpt_family="Rf:ATREPS|ATREPS An AT-rich repetitive
sequence - a consensus."
complement(17361..17388)
/rpt_family="AT_rich"
17599..17657
/rpt_family="AT_rich"
18967..18992
/rpt_family="(TA)n"
19827..19874
/rpt_family="(TA)n"
20755..20797
/rpt_family="(TAAAAA)n"
complement(20851..20873)
/rpt_family="AT_rich"
21014..24475
/gene="At2g19450"
/notes="synonym: F3p11.5; supported by cDNA:
gi_15450799_gb_AY054480.1"
/join(21014..21564,21649..21729,21821..21894,21997..>22072,
22170..22237,22319..22442,22523..22600,22679..>22780,
22881..23017,23162..23243,23348..23368,23453..>23515,
23599..23646,23770..23832,23914..23964,24090..>24475)
/gene="At2g19450"
/join(21175..21564,21649..21729,21821..21894,21997..>22072,
22170..22237,22319..22442,22523..22600,22679..>22780,
22881..23017,23162..23243,23348..23368,23453..>23515,
23599..23646,23770..23832,23914..23964,24090..>24475)
/gene="At2g19450"
/codon_start=1
/product="diacylglycerol O-acyltransferase"
/protein_id="AAD10144.2"
/db_xref="GI:20197480"
/translation="MAILDSAGVTVTVTENGGEFVDLDRRRKSRSDSSNGLLLSGS
DNNPSDDVADPADVRDRIDSVNDDAOGTANLAGDNGGNGGGRGGGRGNAD
ATPYRSPVPAHRRARESPSLDAIFPKSHAGLFNLVVVLAVNSRLIENLMKYGM
LRTDFWFSRSLRWPLFMCCISLIFPLAAFTVEKVLQKIITSEPVVIFLIITGM
TEVLYPVYVTRCDSAFSLGVTMLLTCTVMLKLVSYAHSYDIIRSLANAADKANPEV
SYVLSGLAYFMVAPTLCYQFSPRSACIRKGVARQFAKLVITFGMFTIEQVIN
PIVRNSKPLKGLDLYAIERVLKSVPLNYVLMCMFYCFPKLWLNLIABLLCFGRREF
YDWNNAKSGVDYWRMMNPVHKVMVRHYLPCLRSKTPKTLAITIAFVSAVFHEL
IAPVCRPLKLNALFGLGIMPQVPLVFTNVLQERFGSTVGNMIFWIFCIPGQPMCVLLY
THLMNRKGSMS"
25114..25134
/rpt_family="AT_rich"
complement(25990..26015)
/rpt_family="AT_rich"
26221..26844
/gene="At2g19460"
/notes="synonym: F3p11.6; predicted by genefinder;
supported by cDNA: gi_14334633_gb_AY034990.1"
26221..26844
/gene="At2g19460"
26320..26670
/gene="At2g19460"
/codon_start=1
/product="expressed protein"
/protein_id="AAD10145.1"
/db_xref="GI:4191776"
/translation="MEDYRSRSYDGRSTDLQOYSAHRSDGPDFSGNGMDLRSYS
TSYDTPRIPEQNPQKRRSSSSSWGFVDFDLQRRKRVSYRAYTVEGKUKGSFRK
SFKWKDKCNKLLN"
complement(26742..26772)
/rpt_family="(CA)n"
complement(28356..31273)
/gene="At2g19470"
/notes="synonym: F3p11.7; supported by full length cDNA:
Ceres:2312"

```

```

QY 21 SerAlaMetIleProGluIleAsp 28
Db 1388 TCAACGATGACACCGGACGAAGAT 1411
RESULT 16
AX339706 AX339706 2991 bp DNA linear PAT 10-JAN-2002
LOCUS Sequence 3 from Patent WO0196580.
DEFINITION AX339706
ACCESSION AX339706
VERSION AX339706.1 GI:18135704
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Schmulling, T. and Werner, T.
TITLE Method for modifying plant morphology, biochemistry and physiology
JOURNAL Patent: WO 0196580-A 3 20-DEC-2001;
Schmulling, Thomas (DE); Werner, Tomas (DE)
FEATURES
LOCATION/Qualifiers
source 1..2991
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Alignment Scores:
Pred. No.: 1 58e-05 Length: 2991
Score: 96.50 Matches: 27
Percent Similarity: 48.21% Conservative: 0
Best Local Similarity: 48.21% Mismatches: 1
Query Match: 65.65% Indels: 29
DB: 6 Gaps: 1

US-10-014-101B-32 (1-28) x AX339706 (1-2991)

QY 1 SerAlaSerGlyLeuAlaLeuLeuTyrProThrAsnArgAsnLys----- 15
Db 2573 TCAGCTTCGGGACTCGCTCTTCTATCCAAACCGGAATAA-GTACATATCTTCTT 2631
QY 15 ----- 15
Db 2632 CATTCAATATTATCTTCAAGAACCAAGTAATAATTTCTATGAAGTATGCTGTT 2691
QY 16 -----TTPAspAsnArgMetSerAlaMetIleProGluIleAsp 28
Db 2692 ATTGTTAGATGGACATCGTATGTCGGCGATGATACAGATCGAT 2739

RESULT 17
AX785054 AX785054 2991 bp DNA linear PAT 17-JUL-2003
LOCUS Sequence 3 from Patent WO03050287.
DEFINITION AX785054
ACCESSION AX785054
VERSION AX785054.1 GI:32952890
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Schmulling, T. and Werner, T.
TITLE Method for modifying plant morphology, biochemistry and physiology
JOURNAL Patent: WO 03050287-A 3 19-JUN-2003;
Schmulling, Thomas (DE); Werner, Tomas (DE)
FEATURES
LOCATION/Qualifiers
source 1..2991
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Alignment Scores:
Pred. No.: 1 58e-05 Length: 2991
Score: 96.50 Matches: 27
Percent Similarity: 48.21% Conservative: 0
Best Local Similarity: 48.21% Mismatches: 1
Query Match: 65.65% Indels: 29
DB: 6 Gaps: 1

US-10-014-101B-32 (1-28) x AX339706 (1-2991)

QY 1 SerAlaSerGlyLeuAlaLeuLeuTyrProThrAsnArgAsnLys----- 15
Db 2573 TCAGCTTCGGGACTCGCTCTTCTATCCAAACCGGAATAA-GTACATATCTTCTT 2631
QY 15 ----- 15
Db 2632 CATTCAATATTATCTTCAAGAACCAAGTAATAATTTCTATGAAGTATGCTGTT 2691
QY 16 -----TTPAspAsnArgMetSerAlaMetIleProGluIleAsp 28
Db 2692 ATTGTTAGATGGACATCGTATGTCGGCGATGATACAGATCGAT 2739

RESULT 18
AC005917 AC005917 92822 bp DNA linear PLN 27-FEB-2002
LOCUS Arabidopsis thaliana chromosome 2 clone F3P11 map CIC06E08,
DEFINITION complete sequence.
ACCESSION AC005917
VERSION AC005917.3 GI:20197478
KEYWORDS HTG.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 92822)
AUTHORS Lin, X., Kaul, S., Shea, T. P., Fujii, C. Y., Shen, M., VanAken, S. E.,
Barnstead, M. E., Mason, T. M., Bowman, C. L., Ronning, C. M.,
Benito, M. -I., Carretero, A. J., Creasy, T. H., Buell, C. R., Town, C. D.,
Nierman, W. C., Fraser, C. M. and Venter, J. C.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 92822)
AUTHORS Lin, X.
TITLE Direct Submission
JOURNAL Submitted (09-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
REFERENCE 3 (bases 1 to 92822)
AUTHORS Town, C. D. and Kaul, S.
TITLE Direct Submission
JOURNAL Submitted (27-FEB-2002) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA, cdtown@tigr.org
COMMENT On Apr 18, 2002 this sequence version replaced gi:6598497.
FEATURES
source 1..92822
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="2"
/map="CIC06E08"
/clone="F3P11"
1..27852
/notes="overlap with BAC clone F27F23
(AC003058:69644..97495)."
repeat_region complement(2151..2171)
repeat_region /rpt_family="AT_rich"
repeat_region complement(5234..5268)
repeat_region /rpt_family="AT_rich"
5488..5517
/rpt_family="AT_rich"
6440..8622
gene

```





SGLLALYPTNRKNDNRMSAMIPEDIDVYIIIGLIQSATPKDLPEVSNKRIIFC  
KDSGIKI KOYLHYHTSKEDWIEHFGSKWDFFSKRDLDPKLLSPQDIF"  
1561. 1687  
/gene="Ac2g19500"

## ORIGIN

## Alignment Scores:

Pred. No.: 3e-14 Length: 1687  
Score: 147.00 Matches: 28  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0

US-10-014-101B-32 (1-28) x BT004107 (1-1687)

Qy 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20  
Db 1225 TCAGCTTCGGGACTCGCTCTCTATCCAAACACCGGAATAAATGGACACATCGTATG 1284

Qy 21 SerAlaMetIleProGluIleAsp 28

Db 1285 TCGGCGATGATACAGATCGAT 1308

## RESULT 9

AX339731 AX339731 1575 bp DNA linear PAT 10-JAN-2002  
LOCUS Sequence 28 from Patent WO0196580.

DEFINITION AX339731

ACCESSION AX339731

VERSION AX339731.1 GI:18135724

KEYWORDS

SOURCE Arabidopsis thaliana (thale cress)

## ORGANISM

Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

## REFERENCE

1 Schmulling, T. and Werner, T.  
Method for modifying plant morphology, biochemistry and physiology  
Patent: WO 0196580-A 28 20-DEC-2001;

Schmulling, Thomas (DE); Werner, Tomas (DE)

## FEATURES

Location/Qualifiers

1..1575

/organism="Arabidopsis thaliana"

/mol\_type="unassigned DNA"

/db\_xref="taxon:3702"

## ORIGIN

## Alignment Scores:

Pred. No.: 9.04e-08 Length: 1575  
Score: 108.00 Matches: 19  
Percent Similarity: 78.57% Conservative: 3  
Best Local Similarity: 67.86% Mismatches: 6  
Query Match: 73.47% Indels: 0  
DB: 6 Gaps: 0

US-10-014-101B-32 (1-28) x AX339731 (1-1575)

Qy 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20  
Db 1240 TCAACTTCGGTGTTACTCTCTATCCCAACCGGAATAAATGGACACCGCATG 1299

Qy 21 SerAlaMetIleProGluIleAsp 28

Db 1300 TCAACGATGACACCGGACGAGAT 1323

## RESULT 10

AX507394 AX507394 1575 bp DNA linear PAT 27-SEP-2002  
LOCUS Sequence 2089 from Patent WO0216655.

## DEFINITION

AX507394

ACCESSION AX507394

VERSION AX507394.1 GI:23388631

KEYWORDS

SOURCE Arabidopsis thaliana (thale cress)

## ORGANISM

Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.

Annotation based on July 2002 version of the Arabidopsis genome submitted to Genbank.

#### FEATURES

source  
Location/Qualifiers  
1. .1537  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="2"  
/clone="U20989"  
/ecotype="Columbia"  
/note="This clone is in pUNI 51."  
1. .1537  
/gene="At2g19500"  
1. .1506  
/gene="At2g19500"  
/codon\_start=1  
/evidence="experimental"  
/product="putative cytokinin oxidase"  
/protein\_id="AA064073.1"  
/db\_xref="GI:28973497"  
/translation="MANLRMTLITVLMITKSSNGIKIDLPKSLNLTSTDPSTISA  
ASHDFGNITVPGVICPSSTADISRLQYAANGKSTFQVAAAGQGHSLNGQASVSG  
GVIVNMTCTIDVVSKDKKYADVAAGTLVDVLTAKGVSPVSWTDLHTVCGTL  
SNGGIGGVFRNGPLVSNVLELDVITGKEMLTCSRQNLNPELFGVGLGQFGIITR  
ARIVLDHAPKRAKWFRLMYSDFTTTKQERLISMANDIGVDYLEGQIFLSNGVVDTS  
FPPPSDQSKADLVKHQGIIVYLEVAKYDDPNLPIISKVIDTLTKLSYLPFGFISMH  
DVAYDFDLNRVBNKRSGLGLWELPHPLNLYPKSRILDFHNGVVKDILLKQKSA  
SGLALLYPTNRKNDRNSAMIPEDIDVYIIGLQATPKDLPEVSVNEKIRFC  
KDSGIKIKQYLMHYTSKEDIWIEHFGSKWDFDKRDLFPDKLLSPGDIP"  
1507. .1537  
/gene="At2g19500"

#### 3' UTR

#### ORIGIN

Alignment Scores:  
Pred. No.: 2,7e-14 Length: 1537  
Score: 147.00 Matches: 28  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0  
US-10-014-101b-32 (1-28) x BT005653 (1-1537)  
QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20  
DB 1171 TCAGCTTCGGGACTCGCTCTCTCTATCCACCAACCGAATAAATGGACAATCGTATG 1230  
QY 21 SerAlaMetIleProGluIleAsp 28  
DB 1231 TCGGCGATGATACACAGATCGAT 1254  
RESULT 8  
BT004107  
LOCUS  
DEFINITION  
Arabidopsis thaliana clone RAF15-29-H04 (R20989) putative  
cytokinin oxidase (At2g19500) mRNA, complete cds.  
ACCESSION  
BT004107  
VERSION  
BT004107.1 GI:28393415  
KEYWORDS  
FLI CDNA.  
SOURCE  
Arabidopsis thaliana (thale cress)

#### ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

#### REFERENCE

1 (bases 1 to 1687)  
Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

Arabidopsis Full Length cDNA Clones

#### TITLE

Unpublished

#### JOURNAL

#### REFERENCE

#### AUTHORS

2 (bases 1 to 1687)  
Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

#### TITLE

Direct Submission

#### JOURNAL

#### COMMENT

Submitted (14-FEB-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA  
RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.

Annotation based on July 2002 version of the Arabidopsis genome submitted to Genbank.

#### FEATURES

##### source

Location/Qualifiers  
1. .1687  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="2"  
/clone="RAFL15-29-H04 (R20989)"  
/ecotype="Columbia"  
/note="This clone is in a modified pBluescript vector2 (lambda PS) as a BamHI/XhoI insert."

##### gene

1. .1687  
/gene="At2g19500"

##### 5' UTR

1. .54  
/gene="At2g19500"

##### CDS

55. .1560  
/gene="At2g19500"  
/codon\_start=1  
/evidence="experimental"  
/product="putative cytokinin oxidase"  
/protein\_id="AA042130.1"  
/db\_xref="GI:28393416"

/translation="MANLRMTLITVLMITKSSNGIKIDLPKSLNLTSTDPSTISA  
ASHDFGNITVPGVICPSSTADISRLQYAANGKSTFQVAAAGQGHSLNGQASVSG  
GVIVNMTCTIDVVSKDKKYADVAAGTLVDVLTAKGVSPVSWTDLHTVCGTL  
SNGGIGGVFRNGPLVSNVLELDVITGKEMLTCSRQNLNPELFGVGLGQFGIITR  
ARIVLDHAPKRAKWFRLMYSDFTTTKQERLISMANDIGVDYLEGQIFLSNGVVDTS  
FPPPSDQSKADLVKHQGIIVYLEVAKYDDPNLPIISKVIDTLTKLSYLPFGFISMH  
DVAYDFDLNRVBNKRSGLGLWELPHPLNLYPKSRILDFHNGVVKDILLKQKSA

```

KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE
AUTHORS Schumling, T. and Werner, T.
TITLE Method for modifying plant morphology, biochemistry and physiology
JOURNAL Patent: WO 03050287-A 26 19-JUN-2003;
FEATURES
source Schumling, Thomas (DE) ; Werner, Tomas (DE)
Location/Qualifiers
1. .1506
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Alignment Scores:
Pred. No.: 2.63e-14 Length: 1506
Score: 147.00 Matches: 28
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-014-101B-32 (1-28) x AX785077 (1-1506)
QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAenArgAsnLysTrpAspAsnArgMet 20
DB 1171 TCAGCTTCGGGATCGCTCTCTCTATCCAAACACCGGAATAATGGGACAAATCGTATG 1230
QY 21 SerAlaMetIleProGluIleAsp 28
DB 1231 TCGGCGATGATACAGATCGAT 1254

RESULT 6
AF303978 1506 bp mRNA linear PLN 08-NOV-2000
LOCUS Arabidopsis thaliana cytokinin oxidase (CKX2) mRNA, complete cds.
DEFINITION
ACCESSION AF303978
VERSION AF303978.1 GI:11120507
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 1506)
AUTHORS Bilyeu, K.D., Laskey, J.G., Riekhof, W.R., VanVickle, S. and Morris, R.O.
TITLE A family of cytokinin oxidases from Arabidopsis thaliana
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1506)
AUTHORS Bilyeu, K.D., Laskey, J.G., Riekhof, W.R., VanVickle, S. and Morris, R.O.
TITLE Direct Submission
JOURNAL Submitted (11-SEP-2000) University of Missouri, 210 Waters Hall, Columbia, MO 65211, USA
FEATURES
source Location/Qualifiers
1. .1506
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosomes="II"
1. .1506
/gene="CKX2"
1. .1506
/gene="CKX2"
/codon_start=1
/product="cytokinin oxidase"
/protein_id="AAG30905.1"
/db_xref="GI:11120508"

KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE
AUTHORS Schumling, T. and Werner, T.
TITLE Method for modifying plant morphology, biochemistry and physiology
JOURNAL Patent: WO 03050287-A 26 19-JUN-2003;
FEATURES
source Schumling, Thomas (DE) ; Werner, Tomas (DE)
Location/Qualifiers
1. .1506
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Alignment Scores:
Pred. No.: 2.63e-14 Length: 1506
Score: 147.00 Matches: 28
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-014-101B-32 (1-28) x AF303978 (1-1506)
QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAenArgAsnLysTrpAspAsnArgMet 20
DB 1171 TCAGCTTCGGGATCGCTCTCTCTATCCAAACACCGGAATAATGGGACAAATCGTATG 1230
QY 21 SerAlaMetIleProGluIleAsp 28
DB 1231 TCGGCGATGATACAGATCGAT 1254

RESULT 7
BT005653 1537 bp mRNA linear PLN 15-MAR-2003
LOCUS Arabidopsis thaliana clone U20989 putative cytokinin oxidase
DEFINITION
ACCESSION BT005653
VERSION BT005653.1 GI:28973496
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 1537)
AUTHORS Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
TITLE Arabidopsis Open Reading Frame (ORF) Clones
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1537)
AUTHORS Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
TITLE Direct Submission
JOURNAL Submitted (15-MAR-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
COMMENT The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

```

```

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-014-101B-32 (1-28) x AX339734 (1-84)

QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20
Db 1 TCAGCTTCGGGACTCGCTCTTCTATCCAAACCGGAATAAATGGGACATCGTATG 60

QY 21 SerAlaMetIleProGluIleAsp 28
Db 61 TCGGCGATGATACACAGATCGAT 84

RESULT 2
AX785082
LOCUS AX785082 84 bp DNA linear PAT 17-JUL-2003
DEFINITION Sequence 31 from Patent WO03050287.
ACCESSION AX785082
VERSION AX785082.1 GI:32952913
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Schmullling, T. and Werner, T.
TITLE Method for modifying plant morphology, biochemistry and physiology
JOURNAL Patent: WO 03050287-A 31 19-JUN-2003;
Schmullling, Thomas (DE); Werner, Tomas (DE)
FEATURES
source
Location/Qualifiers
1..84
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Alignment Scores:
Pred. No.: 9,41e-16 Length: 84
Score: 147.00 Matches: 28
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-014-101B-32 (1-28) x AX785082 (1-84)

QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20
Db 1 TCAGCTTCGGGACTCGCTCTTCTATCCAAACCGGAATAAATGGGACATCGTATG 60

QY 21 SerAlaMetIleProGluIleAsp 28
Db 61 TCGGCGATGATACACAGATCGAT 84

RESULT 3
CQ812637
LOCUS CQ812637 1506 bp DNA linear PAT 24-MAY-2004
DEFINITION Sequence 3 from Patent WO2004036027.
ACCESSION CQ812637
VERSION CQ812637.1 GI:47602086
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS van Camp, W.
TITLE Bioremediation
JOURNAL Patent: WO 2004036027-A 3 06-MAY-2004;
CropDesign N.V. (BE)

```

---

```

FEATURES
source
Location/Qualifiers
1..1506
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Alignment Scores:
Pred. No.: 2,63e-14 Length: 1506
Score: 147.00 Matches: 28
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-014-101B-32 (1-28) x CQ812637 (1-1506)

QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20
Db 1171 TCAGCTTCGGGACTCGCTCTTCTATCCAAACCGGAATAAATGGGACATCGTATG 1230

QY 21 SerAlaMetIleProGluIleAsp 28
Db 1231 TCGGCGATGATACACAGATCGAT 1254

RESULT 4
AX339729
LOCUS AX339729 1506 bp DNA linear PAT 10-JAN-2002
DEFINITION Sequence 26 from Patent WO0196580.
ACCESSION AX339729
VERSION AX339729.1 GI:18135722
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Schmullling, T. and Werner, T.
TITLE Method for modifying plant morphology, biochemistry and physiology
JOURNAL Patent: WO 0196580-A 26 20-DEC-2001;
Schmullling, Thomas (DE); Werner, Tomas (DE)
FEATURES
source
Location/Qualifiers
1..1506
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Alignment Scores:
Pred. No.: 2,63e-14 Length: 1506
Score: 147.00 Matches: 28
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-014-101B-32 (1-28) x AX339729 (1-1506)

QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20
Db 1171 TCAGCTTCGGGACTCGCTCTTCTATCCAAACCGGAATAAATGGGACATCGTATG 1230

QY 21 SerAlaMetIleProGluIleAsp 28
Db 1231 TCGGCGATGATACACAGATCGAT 1254

RESULT 5
AX785077
LOCUS AX785077 1506 bp DNA linear PAT 17-JUL-2003
DEFINITION Sequence 26 from Patent WO03050287.
ACCESSION AX785077
VERSION AX785077.1 GI:32952908

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: February 18, 2005, 07:19:31 ; Search time 6601 Seconds  
(without alignments)  
205.536 Million cell updates/sec

Title: US-10-014-101B-32

Perfect score: 147  
Sequence: 1 SASGLALLYPTNRKNDNRMSAMPEID 28

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Deiop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10014101/runat\_16022005\_075934\_8298/app\_query.fasta\_1.199  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0\_1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -ALIGN=40 -MODE=LOCAL  
-OUTFMT=ptc -NOR=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10014101 @CGN\_1\_1\_2527 @runat\_16022005\_075934\_8298 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*

1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	147	100.0	84	AX339734	Sequence
2	147	100.0	84	AX785082	Sequence
3	147	100.0	1506	CQ812637	Sequence
4	147	100.0	1506	AX339729	Sequence

5	147	100.0	1506	6	AX785077	Sequence
6	147	100.0	1506	8	AF303978	Arabidops
7	147	100.0	1537	8	BT005653	Arabidops
8	147	100.0	1687	8	BT004107	Arabidops
9	108	73.5	1575	6	AX339731	Sequence
10	108	73.5	1575	6	AX507394	Sequence
11	108	73.5	1575	6	AX651615	Sequence
12	108	73.5	1575	6	AX785079	Sequence
13	108	73.5	1575	8	AF303980	Arabidops
14	108	73.5	1655	8	BT000179	Arabidops
15	108	73.5	1863	8	AY054460	Arabidops
16	96.5	65.6	2291	6	AX339706	Sequence
17	96.5	65.6	2291	6	AX785054	Sequence
18	96.5	65.6	2282	8	AC005917	Arabidops
19	96	65.3	1572	6	AX339730	Sequence
20	96	65.3	1572	6	AX785078	Sequence
21	96	65.3	1572	8	AF303979	Arabidops
22	89	60.5	2134	8	AF303979	Arabidops
23	88	59.9	1590	6	AX653226	Sequence
24	88	59.9	2328	8	AK121317	Sequence
25	87	59.2	1677	6	AX653630	Sequence
26	86	58.5	2146	8	AF303982	Arabidops
27	85	57.8	1587	6	AX653027	Sequence
28	83	56.5	1566	6	AX654568	Sequence
29	83	56.5	2264	8	AF303979	Arabidops
30	83	56.5	2782	6	AX339710	Sequence
31	83	56.5	2782	6	AX785058	Sequence
32	83	56.5	8801	8	AT116144	Arabidops
33	83	56.5	200001	8	ATCHRIV71	Arabidops
34	81	55.1	1930	8	AK101022	Oryza sat
35	80	54.4	1515	6	AX339733	Sequence
36	80	54.4	1515	6	AX785081	Sequence
37	80	54.4	1740	8	DSO294542	Dendrobii
38	79	53.7	1866	8	AY209184	Hordeum v
39	77	52.4	342	11	BX255834	Arabidops
40	77	52.4	366	11	BX255816	Arabidops
41	77	52.4	1611	6	AX339732	Sequence
42	77	52.4	1611	6	AX785080	Sequence
43	77	52.4	1620	6	AX339737	Sequence
44	77	52.4	1620	6	AX785085	Sequence
45	77	52.4	1623	8	AF303982	Arabidops

#### ALIGNMENTS

AX339734	Sequence 31 from Patent WO0196580.	84 bp	DNA	linear	PAT 10-JAN-2002
AX339734	AX339734				
AX339734.1	GI:18135727				
Arabidopsis thaliana (thale cress)					
Arabidopsis thaliana					
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.					
Schmueling, T. and Werner, T.					
Method for modifying plant morphology, biochemistry and physiology					
Patent: WO 0196580-A 31 20-DEC-2001;					
Schmueling, Thomas (DE); Werner, Tomas (DE)					
Location/Qualifiers					
1..84					
/organism="Arabidopsis thaliana"					
/mol_type="unassigned DNA"					
/db_xref="taxon:3702"					

#### ORIGIN

Alignment Scores:  
Pred. No.: 9.41e-16 Length: 84  
Score: 147.00 Matches: 28  
Percent Similarity: 100.00% Conservative: 0

**This Page Blank (uspto)**

```

RN      [1]
RP      SEQUENCE FROM N.A.
RA      Bilyeu K.D., Laskey J.G., Riekhof W.R., VanVickle S., Morris R.O.;
RL      Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RR      EMBL; AF303982; AAC30909.1; -.
DR      GO; GO:0006118; P:electron transport; IEA.
DR      InterPro; IPR006094; Oxid_FAD_bind_N.
DR      InterPro; IPR006093; Oxred_FAD_bind.
DR      Pfam; PF01585; FAD_binding_4; 1.
DR      PROSITE; PS00862; OX2_COVAL_FAD; UNKNOWN 1.
DR      PROSITE; PS00862; OX2_COVAL_FAD; UNKNOWN 1.
SQ      SEQUENCE 540 AA; 60389 MW; D74DD0D50CEBA76F CRC64;

Query Match          52.4%; Score 77; DB 2; Length 540;
Best Local Similarity 54.2%; Pred.No. 0.0069;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy      3  SGLALLYPTNRNKWDNRMSAMPE 26
         ||| :||| :||| :||| :
Db      409  SGPILIIYPMNKRWDERSAVTPD 432

```

Search completed: February 22, 2005, 22:44:15  
Job time : 179 secs

Query Match 54.4%; Score 80; DB 2; Length 536;  
Best Local Similarity 60.9%; Pred. No. 0.0024;  
Matches 14; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 4 GLALLYPTRNKNWDRMSAMIPE 26  
DB 419 GILVLPYTKRSKWDKRMSTSIPT 441

## RESULT 12

ID Q84U27 PRELIMINARY; PRT; 520 AA.  
AC Q84U27;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Cytochrome dehydrogenase 3 (EC 1.5.99.12).  
OS Hordeum vulgare (Barley).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
OC Triticeae; Hordeum.  
OC NCBI\_TaxID=4513;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Galuszka P., Frebort I.;  
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY209184; AA050082.1; -;  
DR GO; GO:0019139; P:cytochrome dehydrogenase activity; IEA.  
DR GO; GO:0016491; P:oxidoreductase activity; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR InterPro; IPR006094; Oxid\_FAD\_bind\_N.  
DR InterPro; IPR006093; Oxid\_FAD\_BS.  
DR Pfam; PF01565; FAD binding\_4; 1.  
DR PROSITE; PS00862; OX2\_COVAL\_FAD; UNKNOWN\_1.  
KW Oxidoreductase.  
SQ SEQUENCE 520 AA; 58163 MW; A6FE752806458AD0 CRC64;

Query Match 53.7%; Score 79; DB 2; Length 520;  
Best Local Similarity 50.0%; Pred. No. 0.0033;  
Matches 12; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 3 SGLALLYPTNRKNWDRMSAMIPE 26  
DB 402 NGPILLYPNKSRWDRNRTSVLPD 425

## RESULT 13

ID Q9FWT3 PRELIMINARY; PRT; 512 AA.  
AC Q9FWT3;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE F1B16.2 protein.  
GN Name=F1B16.2;  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurooids II; Brassicales; Brassicaceae; Arabidopsis.  
OC NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Federpsiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,  
RA Altafi H., Nguyen M., Lam B., Southwick A., Miranda M., Brooks S.,  
RA Buehler E., Chao O., Chin C., Chioi J., Choi E., Gonzalez A.,  
RA Hwang B., Johnson-Hopson C., Khan S., Kim C., Koo T., Lee J.M.,  
RA Lenz C., Liu A., Liu S., Mukharsky N., Pham P., Sakano H., Shinn P.,  
RA Toriumi M., Vaysberg M., Yu G., Ecker J., Theologis A., Davis R.W.;  
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC023754; AAG13068.1; -;  
DR PIR; B96785; B96785.  
DR GO; GO:0005622; C:intracellular; IEA.  
DR GO; GO:0005640; C:ribosome; IEA.

DR GO; GO:0003735; F:structural constituent of ribosome; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR GO; GO:0006412; P:protein biosynthesis; IEA.  
DR InterPro; IPR006094; Oxid\_FAD\_bind\_N.  
DR InterPro; IPR006093; Oxid\_FAD\_BS.  
DR InterPro; IPR002132; Ribosomal\_L5.  
DR Pfam; PF01565; FAD\_binding\_4; 1.  
DR PROSITE; PS00862; OX2\_COVAL\_FAD; UNKNOWN\_1.  
DR PROSITE; PS00358; RIBOSOMAL\_L5; 1.  
SQ SEQUENCE 512 AA; 57292 MW; 4CA5B3F93EEFDA51 CRC64;

Query Match 52.4%; Score 77; DB 2; Length 512;  
Best Local Similarity 54.2%; Pred. No. 0.0065;  
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 3 SGLALLYPTNRKNWDRMSAMIPE 26  
DB 381 SGPILLYPMNKDKWDRSSAVTPD 404

## RESULT 14

ID Q67YU0 PRELIMINARY; PRT; 540 AA.  
AC Q67YU0;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Cytochrome oxidase (CKX6).  
GN Name=At1g75450;  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurooids II; Brassicales; Brassicaceae; Arabidopsis.  
OC NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Totoki Y., Seki M., Ishida J., Nakajima M., Enju A., Kamiya A.,  
RA Narusaka M., Shin-i T., Nakagawa M., Sakamoto N., Oishi K., Kohara Y.,  
RA Kobayashi M., Toyoda A., Sakaki Y., Sakurai T., Iida K., Akiyama K.,  
RA Satou M., Toyoda T., Konagaya A., Carninci P., Kawai J.,  
RA Hayashizaki Y., Shinozaki K.;  
RT "Large-scale analysis of RIKEN Arabidopsis full-length (RAFL) cDNAs";  
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK176378; BAD4141.1; -;  
DR InterPro; IPR006094; Oxid\_FAD\_bind\_N.  
DR InterPro; IPR006093; Oxid\_FAD\_BS.  
DR Pfam; PF01565; FAD binding\_4; 1.  
DR PROSITE; PS00862; OX2\_COVAL\_FAD; UNKNOWN\_1.  
SQ SEQUENCE 540 AA; 60423 MW; 9B3BA1138E252B84 CRC64;

Query Match 52.4%; Score 77; DB 2; Length 540;  
Best Local Similarity 54.2%; Pred. No. 0.0069;  
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 3 SGLALLYPTNRKNWDRMSAMIPE 26  
DB 409 SGPILLYPMNKDKWDRSSAVTPD 432

## RESULT 15

ID Q9FUJ0 PRELIMINARY; PRT; 540 AA.  
AC Q9FUJ0;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Cytochrome oxidase.  
GN Name=CKX6;  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurooids II; Brassicales; Brassicaceae; Arabidopsis.  
OC NCBI\_TaxID=3702;



ID Q8S0F8 PRELIMINARY; PRT; 525 AA.  
 AC Q8S0F8;  
 DT 01-JUN-2002 (TRENBLrel. 21, Created)  
 DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)  
 DT 01-WAR-2004 (TRENBLrel. 26, Last annotation update)  
 DE Putative cytokinin oxidase.  
 GN Name=B1150fill119;  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=12447438; DOI=10.1038/nature01184;  
 RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,  
 RA Wu J., Nilimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,  
 RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,  
 RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,  
 RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,  
 RA Ikono M., Itoh S., Itoh T., Itoh Y., Iwabuchi A., Kamiya K.,  
 RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,  
 RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,  
 RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,  
 RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,  
 RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,  
 RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,  
 RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,  
 RA Yano M., Jiang J., Gojobori T.;  
 RT "The genome sequence and structure of rice chromosome 1.";  
 RL Nature 420:312-316(2002).  
 DR EMBL; AP003412; BAB90259.1; -;  
 DR Gramene; Q8S0F8; -;  
 DR InterPro; IPR006094; Oxid\_FAD\_bind\_N.  
 DR InterPro; IPR006093; Oxred\_FAD\_BS.  
 DR Pfam; PF01565; FAD\_binding\_4; 1.  
 DR PROSITE; PS00862; OX2\_COVAL\_FAD; UNKNOWN 1.  
 SQ SEQUENCE 525 AA; 58011 MW; D10B03BFD4259DDD CRC64;  
  
 Query Match 59.9%; Score 88; DB 2; Length 525;  
 Best Local Similarity 61.5%; Pred. No. 0.00015;  
 Matches 16; Conservative 3; Mismatches 7; Indels 0; Gaps 0;  
  
 QY 1 SASGLALLYPTNRNKNWDRMSAMPEID 26  
 Db 404 SNNGPILLYPVNRKWNDRNRTSVVDP 429  
  
 RESULT 6  
 Q941V9  
 ID Q941V9 PRELIMINARY; PRT; 558 AA.  
 AC Q941V9;  
 DT 01-DEC-2001 (TRENBLrel. 19, Created)  
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)  
 DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)  
 DE Putative cytokinin oxidase.  
 GN Name=P0419B01.11; Synonyms=B1046G12.5;  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=12447438; DOI=10.1038/nature01184;  
 RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,  
 RA Wu J., Nilimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,  
 RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,  
 RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,  
 RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,  
 RA Ikono M., Itoh S., Itoh T., Itoh Y., Iwabuchi A., Kamiya K.,  
 RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,  
 RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,  
 RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,  
 RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,  
 RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,  
 RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,  
 RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,  
 RA Yano M., Jiang J., Gojobori T.;  
 RT "The genome sequence and structure of rice chromosome 1.";  
 RL Nature 420:312-316(2002).  
 DR EMBL; AP003412; BAB90259.1; -;  
 DR Gramene; Q8S0F8; -;  
 DR InterPro; IPR006094; Oxid\_FAD\_bind\_N.  
 DR InterPro; IPR006093; Oxred\_FAD\_BS.  
 DR Pfam; PF01565; FAD\_binding\_4; 1.  
 DR PROSITE; PS00862; OX2\_COVAL\_FAD; UNKNOWN 1.  
 SQ SEQUENCE 525 AA; 58011 MW; D10B03BFD4259DDD CRC64;  
  
 Query Match 59.9%; Score 88; DB 2; Length 525;  
 Best Local Similarity 61.5%; Pred. No. 0.00015;  
 Matches 16; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,  
 RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,  
 RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,  
 RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,  
 RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,  
 RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,  
 RA Yano M., Jiang J., Gojobori T.;  
 RT "The genome sequence and structure of rice chromosome 1.";  
 RL Nature 420:312-316(2002).  
 DR EMBL; AP003244; BAB56095.1; -;  
 DR EMBL; AP003200; BAB89407.1; -;  
 DR Gramene; Q941V9; -;  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR006094; Oxid\_FAD\_bind\_N.  
 DR Pfam; PF01565; FAD\_binding\_4; 1.  
 SQ SEQUENCE 558 AA; 59121 MW; B3472B591AD1EFF7 CRC64;  
  
 Query Match 59.2%; Score 87; DB 2; Length 558;  
 Best Local Similarity 59.3%; Pred. No. 0.00022;  
 Matches 16; Conservative 4; Mismatches 7; Indels 0; Gaps 0;  
  
 QY 2 ASGLALLYPTNRNKNWDRMSAMPEID 28  
 Db 440 AMGPVLLYPMNRNKNWDSNMSAVITDDD 466  
  
 RESULT 7  
 Q709Q3  
 ID Q709Q3 PRELIMINARY; PRT; 525 AA.  
 AC Q709Q3;  
 DT 05-JUL-2004 (TRENBLrel. 27, Created)  
 DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)  
 DE Cytokinin oxidase 3.  
 GN Name=cko3;  
 OS Zea mays (Maize).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC PACCAD clade; Panicoideae; Andropogoneae; Zea.  
 OX NCBI\_TaxID=4577;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kernal;  
 RA Massonneau A., Houba-Herlin N., Pethe C., Madzak C., Majira A.,  
 RA Falque M., Rogowsky P., Laloue M.;  
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ606944; CAES5202.1; -;  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR006094; Oxid\_FAD\_bind\_N.  
 DR InterPro; IPR006093; Oxred\_FAD\_BS.  
 DR Pfam; PF01565; FAD\_binding\_4; 1.  
 DR PROSITE; PS00862; OX2\_COVAL\_FAD; UNKNOWN 1.  
 SQ SEQUENCE 525 AA; 58504 MW; 56A43B6B255ECD6E CRC64;  
  
 Query Match 58.5%; Score 86; DB 2; Length 525;  
 Best Local Similarity 57.7%; Pred. No. 0.00029;  
 Matches 15; Conservative 5; Mismatches 6; Indels 0; Gaps 0;  
  
 QY 1 SASGLALLYPTNRNKNWDRMSAMPEID 26  
 Db 404 SNNGPILLYPVNRKWNDRNRTSVVDP 429  
  
 RESULT 8  
 Q709Q5  
 ID Q709Q5 PRELIMINARY; PRT; 519 AA.  
 AC Q709Q5;  
 DT 05-JUL-2004 (TRENBLrel. 27, Created)  
 DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)  
 DE Cytokinin oxidase 2.  
 GN Name=cko2;  
 OS Zea mays (Maize).

```

DR EMBL; AL079344; CAB45334.1; --
DR EMBL; AL161575; CAB97932.1; --
DR EMBL; AY054460; AAK96652.1; --
DR EMBL; BT000179; AAN15498.1; --
DR PIR; T09937; T09937.
DR InterPro; IPR006094; Oxid_FAD_bind_N.
DR InterPro; IPR006093; Oxred_FAD_BS.
DR Pfam; PF01565; FAD_binding_4; 1.
DR PROSITE; PS00862; OX2_COVAL_FAD; FALSE NEG.
KW Alternative splicing; FAD; Flavoprotein; Glycoprotein;
KW Multigene family; Oxidoreductase; Signal.
FT SIGNAL 1 26 Potential.
FT CHAIN 27 524 Cytokinin dehydrogenase 4.
FT MOD_RES 109 109 Tele-8alpha-FAD histidine (By
FT similarity).
FT CARBOHYD 39 39 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 58 58 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 124 124 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 411 411 N-linked (GlcNAc... ) (Potential).
FT CONFLICT 15 15 L -> F (in Ref. 1).
SQ SEQUENCE 524 AA; 58061 MW; DBD23A7876DDA171 CRC64;

Query Match 73.5%; Score 108; DB 1; Length 524;
Best Local Similarity 67.9%; Pred. No. 1.4e-07;
Matches 19; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 SASGALLYPTNRNKNDRMSAMIPEID 28
DB 414 STSGVTLFYPTNRNKNDRMSMTPTDED 441

RESULT 3
CKX3_ARATH
ID CKX3_ARATH STANDARD; PRT; 523 AA.
AC Q9LTS3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Cytokinin dehydrogenase 3 precursor (EC 1.5.99.12) (Cytokinin oxidase
DE 3) (CKO 3).
GN Name=CKX3; OrderedLocusNames=At5g56970; ORFNames=MHM17.8;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=21068113; PubMed=11154345; DOI=10.1104/pp.125.1.378;
RA Bilyeu K.D., Cole J.L., Laskey J.G., Riekhof W.R., Esparza T.J.,
RA Kramer M.D., Morris R.O.;
RT "Molecular and biochemical characterization of a cytokinin oxidase
RT from maize."
RL Plant Physiol. 125:378-386(2001).
[2]
SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and PAC
RT clones."
RL DNA Res. 7:31-63(2000).
CC -!- FUNCTION: Catalyzes the oxidation of cytokinins, a family of N(6)-
CC substituted adenine derivatives that are plant hormones, where the
CC substituent is an isopentenyl group.
CC -!- CATALYTIC ACTIVITY: N(6)-dimethylallyladenine + acceptor + H(2)O =
CC adenine + 3-methylbut-2-enal + reduced acceptor.
CC -!- COFACTOR: FAD (By similarity).
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- SIMILARITY: Belongs to the oxygen-dependent FAD-linked
CC oxidoreductase family.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF03979; AAG30906.1; --
DR EMBL; AB024035; BAA97027.1; --
DR InterPro; IPR006094; Oxid_FAD_bind_N.
DR InterPro; IPR006093; Oxred_FAD_BS.
DR Pfam; PF01565; FAD_binding_4; 1.
DR PROSITE; PS00862; OX2_COVAL_FAD; 1.
DR FAD; Flavoprotein; Glycoprotein; Multigene family; Oxidoreductase;
KW Signal.
KW SIGNAL 1 31 Potential.
FT CHAIN 32 523 Cytokinin dehydrogenase 3.
FT MOD_RES 105 105 Tele-8alpha-FAD histidine (By
FT similarity).
FT CARBOHYD 153 153 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 408 408 N-linked (GlcNAc... ) (Potential).
SQ SEQUENCE 523 AA; 59422 MW; 2324EBFC21D7103A CRC64;

Query Match 65.3%; Score 96; DB 1; Length 523;
Best Local Similarity 69.2%; Pred. No. 8.9e-06;
Matches 18; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 SGLALLYPTNRNKNDRMSAMIPEID 28
DB 411 SGPVLVYPNTRNKNDRMSAIPED 436

RESULT 4
Q709Q4
ID Q709Q4 PRELIMINARY; PRT; 525 AA.
AC Q709Q4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cytokinin oxidase 3.
GN Name=cko3;
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
[1]
SEQUENCE FROM N.A.
RC TISSUE=Kernel;
RA Massonneau A., Houbart-Herlin N., Pethe C., Madzak C., Majira A.,
RA Falque M., Rogowsky P., Laloue M.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ606943; CAES5201.1; --
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR006094; Oxid_FAD_bind_N.
DR InterPro; IPR006093; Oxred_FAD_BS.
DR Pfam; PF01565; FAD_binding_4; 1.
DR PROSITE; PS00862; OX2_COVAL_FAD; UNKNOWN 1.
SQ SEQUENCE 525 AA; 58494 MW; 017201BBA5130F49 CRC64;

Query Match 60.5%; Score 89; DB 2; Length 525;
Best Local Similarity 61.5%; Pred. No. 0.0001;
Matches 16; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 SASGALLYPTNRNKNDRMSAMIPE 26
DB 404 SNNGPILLYPNRKNDRNRTSVVIPD 429

RESULT 5
Q8S0F8

```

or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch).

CC -----  
 DR EMBL; AF303978; AAG30905.1; -  
 DR EMBL; AC005917; AAD10149.2; -  
 DR InterPro; IPR006094; Oxid\_FAD\_bind\_N.  
 DR Pfam; PF01565; FAD\_binding\_4; 1.  
 DR PROSITE; PS00862; OX2\_COVAL\_FAD; 1.  
 KW FAD; Flavoprotein; Glycoprotein; Multigene family; Oxidoreductase;  
 Signal.  
 FT SIGNAL 1 22 Potential.  
 FT CHAIN 23 501 Cytokinin dehydrogenase 2.  
 FT MOD\_RES 92 92 Tele-galpa-FAD histidine (By similarity).  
 FT CARBOHYD 32 32 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 51 51 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 107 107 N-linked (GlcNAc...) (Potential).  
 FT CONFLICT 157 157 G -> R (in Ref. 1).  
 SQ SEQUENCE 501 AA; 55583 MW; 9F8F0AAEA4DE84A CRC64;  
 Query Match 100.0%; Score 147; DB 1; Length 501;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-13;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SASGLALYPTNRNKNWDRMSAMIPED 28  
 DB 391 SASGLALYPTNRNKNWDRMSAMIPED 418  
 RESULT 2  
 CKX4\_ARATH STANDARD; PRT; 524 AA.  
 AC Q9FJG2; Q9S077;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Cytokinin dehydrogenase 4 precursor (EC 1.5.99.12) (Cytokinin oxidase 4) (CKO 4).  
 GN Names:CKX4; OrderedLocusNames=At4g23740; ORFNames=T16L4.250;  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopses.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21068113; PubMed=11154345; DOI=10.1104/pp.125.1.378;  
 RA Bilyeu K.D., Cole J.L., Laskey J.G., Riekhof W.R., Esparza T.J.,  
 RA Kramer M.D., Morris R.O.;  
 RA "Molecular and biochemical characterization of a cytokinin oxidase from maize."  
 RT Plant Physiol. 125:378-386(2001).  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=20083488; PubMed=10617198; DOI=10.1038/47134;  
 RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,  
 RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,  
 RA Harris B., Ansoorge W., Brandt P., Grivell L.A., Rieger M.,  
 RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,  
 RA Kreis M., Belseny M., Puigdomenech P., Watson M., Schmidtheini T.,  
 RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,  
 RA Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,  
 RA Langham S.-A., McCullagh B., Bilham J., Robben J.,  
 RA Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,  
 RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,  
 RA Weitzenecker T., Sothe G., Rumpelger U., Hilbert H., Braun M.,  
 RA Holzer B., Brandt A., Peters S., van Staveren M., Dirksse W.,  
 RA Boorjman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,  
 RA Bernisier S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,  
 RA De Keyser A., Buysshaert C., Gielen J., Villarroel R., De Clercq R.,  
 RA Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray-Allen S.,  
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIay K., Mayes R.,

Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,  
 Borikova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,  
 Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,  
 Gabel C., Fuchs M., Partmann B., Granderath K., Dauner D., Herzl A.,  
 Neumann S., Argirou A., Vitale D., Liguori R., Piravandi E.,  
 Massenot O., Quigley F., Clabaud G., Muendlein A., Felber R.,  
 Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,  
 Cherdor F., Cooke R., Berger C., Monfort A., Casacuberta E.,  
 Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,  
 Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,  
 Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C.,  
 Frihman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,  
 Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,  
 Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,  
 Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,  
 Stokeling T., Kalicki J., Graves T., Harmon G., Edwards J.,  
 Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,  
 Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,  
 Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.W.,  
 Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,  
 Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,  
 Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,  
 Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,  
 Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,  
 Granat S., Shondy N., Hasegawa A., Rameed A., Lodhi M., Johnson A.,  
 Chen E., Marra M.A., Martienssen R., McCombie W.R.;  
 RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana."  
 RL Nature 402:769-777(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=22954850; PubMed=14593172; DOI=10.1126/science.1088305;  
 RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,  
 Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,  
 Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,  
 Miranda M., Quach H.B., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,  
 Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,  
 Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,  
 Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,  
 Hayashizaki Y., Johnson-Hopson C., Huan V.W., Iida K., Karnes M.,  
 Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,  
 Kamaya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,  
 Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,  
 Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;  
 RT "Empirical analysis of transcriptional activity in the Arabidopsis genome."  
 RL Science 302:842-846(2003).  
 CC -!- FUNCTION: Catalyzes the oxidation of cytokinins, a family of N(6)-substituted adenine derivatives that are plant hormones, where the substituent is an isopentenyl group.  
 CC -!- CATALYTIC ACTIVITY: N(6)-dimethylallyladenine + acceptor + H(2)O = adenine + 3-methylbut-2-enal + reduced acceptor.  
 CC -!- COFACTOR: FAD (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Extracellular.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=1;  
 CC Comment=A number of isoforms are produced. According to EST sequences;  
 CC Name=1;  
 CC IsoId=Q9FU2-1; Sequence=Displayed;  
 CC -!- SIMILARITY: Belongs to the oxygen-dependent FAD-linked oxidoreductase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL; AF303980; AAG30907.1; -

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 22, 2005, 22:23:01 ; Search time 177 Seconds  
(without alignments)  
81.007 Million cell updates/sec

Title: US-10-014-101B-32

Perfect score: 147

Sequence: 1 SASGALLYPTNRKNKWNRSAMTPEID 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	147	100.0	501	1 CKX2_ARATH	O9fu33 arabidopsis
2	108	73.5	524	1 CKX4_ARATH	O9fu32 arabidopsis
3	96	65.3	523	1 CKX3_ARATH	O9lte3 arabidopsis
4	89	60.5	525	2 Q709Q4	O709Q4 zea mays (m
5	88	59.9	525	2 O8S0F8	O8S0F8 oryza sativ
6	87	59.2	558	2 Q941V9	O941V9 oryza sativ
7	86	58.5	525	2 Q709Q3	O709Q3 zea mays (m
8	83	56.5	519	2 Q709Q5	O709Q5 zea mays (m
9	83	56.5	521	2 Q75K78	O75K78 oryza sativ
10	80	54.4	504	2 Q9LY71	O9LY71 arabidopsis
11	80	54.4	536	2 Q9FE45	O9FE45 dendrobium
12	79	53.7	520	2 Q84U27	O84U27 hordeum vul
13	77	52.4	512	2 Q9FWT3	O9FWT3 arabidopsis
14	77	52.4	540	2 Q67YU0	O67YU0 arabidopsis
15	77	52.4	540	2 Q9FUU0	O9FUU0 arabidopsis
16	75	51.0	526	2 Q8S394	O8S394 hordeum vul
17	75	51.0	526	2 O8H6F6	O8H6F6 hordeum vul
18	75	51.0	527	2 Q6YW51	O6YW51 oryza sativ
19	74	50.3	532	1 CKX1_ORYSA	O9lde6 oryza sativ
20	70	47.6	241	2 Q6QJF4	O6QJF4 fagopyrum e
21	69	46.9	524	2 Q9FUU1	O9FUU1 arabidopsis
22	68	46.3	137	2 Q94K14	O94K14 hordeum vul
23	68	46.3	137	2 Q94K15	O94K15 triticum ae
24	63	46.3	527	2 Q8LNV6	O8LNV6 oryza sativ
25	63	42.9	534	1 CKX1_MAIZE	O9t0n8 zea mays (m
26	62	42.2	524	2 Q6YW50	O6YW50 oryza sativ
27	58	39.5	575	2 O22213	O22213 arabidopsis
28	57	38.8	532	2 Q7XKG2	O7XKG2 oryza sativ
29	56	38.1	222	2 Q6LGM6	O6LGM6 photobacter
30	56	38.1	1274	1 NPC2_CABEL	P34389 caenorhabdi
31	55	37.4	120	2 Q8R9F1	O8R9F1 thermoanaer

32	54	36.7	204	2	Q8U6D2	Q8U6D2 agrobacteri
33	53.5	36.4	668	2	Q06677	Q06677 saccharomyc
34	53	36.1	688	2	Q9HR07	Q9HR07 halobacteri
35	52.5	35.7	227	2	Q6W6I7	Q6W6I7 anseranas s
36	52.5	35.7	242	2	Q8KPD9	Q8KPD9 synchococc
37	52	35.4	269	2	Q77NK3	Q77NK3 rhesus monk
38	52	35.4	269	2	Q9WRS5	Q9WRS5 macaca mula
39	51.5	35.0	206	2	Q6UJC2	Q6UJC2 megadyptes
40	51.5	35.0	227	2	Q8LTZ4	Q8LTZ4 chlorostilb
41	51.5	35.0	370	2	Q937S4	Q937S4 bruceella me
42	51.5	35.0	498	1	GLPK_BRUME	Q937S4 bruceella me
43	51.5	35.0	498	1	GLPK_BRUSU	Q8FVK8 bruceella su
44	51	34.7	518	2	Q6Z9S5	Q6Z9S5 oryza sativ
45	51	34.7	708	2	Q6WFZ6	Q6WFZ6 macaca fasc

## ALIGNMENTS

RESULT 1  
CKX2\_ARATH  
ID CKX2\_ARATH STANDARD; PRT; 501 AA.  
AC Q9FUJ3; Q9ZUP1;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Cytokinin dehydrogenase 2 precursor (EC 1.5.99.12) (Cytokinin oxidase  
DE 2) (CKO 2)  
GN Name=CKX2; OrderedLocustNames=At2g19500; ORFNames=F3P11.10;  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21068113; PubMed=11154345; DOI=10.1104/pp.125.1.378;  
RA Bilyeu K.D., Cole J.L., Laskey J.G., Riekhof W.R., Esparza T.J.,  
RA Kramer M.D., Morris R.O.;  
RT "Molecular and biochemical characterization of a cytokinin oxidase  
RT from maize."  
RL Plant Physiol. 125:378-386(2001).  
[2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RC MEDLINE=20083487; PubMed=10617197; DOI=10.1038/45471;  
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,  
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,  
RA Buell C.R., Ketchum K.A., Lee J.J., Renning C.M., Koo H.L.,  
RA Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,  
RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,  
RA Goodman H.M., Somerville C.R., Coppenhaver G.P., Preuss D.,  
RA Nieman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,  
RA Venter J.C.;  
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
RT thaliana."  
RL Nature 402:761-768(1999).  
CC -!- FUNCTION: Catalyzes the oxidation of cytokinins, a family of N(6)-  
CC substituted adenine derivatives that are plant hormones, where the  
CC substituent is an isopentenyl group.  
CC -!- CATALYTIC ACTIVITY: N(6)-dimethylallylamine + acceptor + H(2)O =  
CC adenine + 3-methylbut-2-enal + reduced acceptor.  
CC -!- COFACTOR: FAD (by similarity).  
CC -!- SUBCELLULAR LOCATION: Extracellular.  
CC -!- SIMILARITY: Belongs to the oxygen-dependent FAD-linked  
CC oxidoreductase family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>)

RESULT 14

H69824

hypothetical protein yhdB - Bacillus subtilis

C/Species: Bacillus subtilis

C/Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004 .

C/Accession: H69824

R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berteaux, C.; Bron, S.; Brouillet, S.; Bruschini, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chao, A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabbet, C.; Ferrari, E.; Nature 390, 249-256, 1997

A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier

```

Query Match      33.3%; Score 49; DB 2; Length 280;
Best Local Similarity 47.1%; Pred. No. 9.7;
Matches 8; Conservative. 5; Mismatches 4; Indels 0; Gaps 0;

QY      2 ASGLALLYPTNRNKNWDN 18
          :|||||:|:|:|:|:
Db      97 SSGIATYPRSRPLWEH 113

```

Search completed: February 22, 2005, 22:45:00  
Job time : 41 secs

A;Cross-references: UNIPROT:Q8U6D2; GB:AEO08689; PIDN:AAL45669.1; PID:g17743395; GSPDB:S59786

C;Experimental source: strain C58 (Dupont)

A;Gene: Atu4875

A;Map position: linear chromosome

Query Match            36.7%; Score 54; DB 2; Length 204;  
Best Local Similarity 46.2%; Pred.No.1.2;  
Matches 12; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 1 SASGLALLYPTNRKWNRMASAMPE 26  
    :|::|||:  
DB 50 SSRGNDLILPSTNNKLRRHMSRE 75

RESULT 10

S59786  
hypothetical protein YDR320c - yeast (*Saccharomyces cerevisiae*)  
N;Alternate names: hypothetical protein D9798.l0  
C;Species: Saccharomycetes cerevisiae  
C;Date: 13-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 09-Jul-2004  
C;Accession: S59786

R;Du, Z.  
submitted to the EMBL Data Library, July 1995

A;Description: The sequence of *S. cerevisiae* cosmid 9798.

A;Reference number: S59418

A;Accession: S59786

A:Molecule type: DNA

A;Residues: 1-668 <DUZ>

A;Cross-references: UNIPROT:QO6677; EMBL:U32517; NID:g914989; PID:g914999; GSPDB:GN0000

C;Genetics:

A;Gene: SGD:YDR320C

A;Cross-references: SGD:S0002728

A;Map position: 4R

Query Match            36.4%; Score 53.5; DB 2; Length 668;  
Best Local Similarity 44.0%; Pred.No.5.3;  
Matches 11; Conservative 7; Mismatches 6; Indels 1; Gaps 1;

QY 3 SGLAL-LYPTRNKWKDNRMASAMPE 26  
    |:||:||||:  
DB 435 SSMALELPSSAKWKNKSNSDP 459

RESULT 11

DB4248  
hypothetical protein Vng0920h [imported] - Halobacterium sp. NRC-1  
C;Species: Halobacterium sp. NRC-1  
C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004

C;Accession: DB4248

R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky,  
J.; Leithauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl,  
Jung, K.H.; Alam, M.; Freitas, T.  
*Proc. Natl. Acad. Sci. U.S.A.* 97, 12176-12181, 2000

A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; L  
A>Title: Genome sequence of Halobacterium species NRC-1.

A;Reference number: RA4160; MUID:20504483; PMID:11016950

A;Accession: DB4248

A>Status: preliminary

A:Molecule type: DNA

A;Residues: 1-668 <STO>

A;Cross-references: UNIPROT:Q9HR07; GB:AE004437; NID:g10580480; PIDN:AAG19352.1; GSPDB

C;Genetics:

A;Gene: VNG0920H

Query Match            36.1%; Score 53; DB 2; Length 688;  
Best Local Similarity 43.5%; Pred.No.6.5;  
Matches 10; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 1 SASGLALLYPTNRKWNRMASAM 23  
    :|::|||:  
DB 109 AGNVALLYGVRDAWDHTAAV 131

Matches 19; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 SASGLALYPTNRKNDNRMSAMPEID 28  
 |||:|||||:|||||:|  
 Db 414 STSGVTLFYPTNRKNNRMSMTWTPED 441

RESULT 3  
 T49185  
 cytochrome oxidase-like protein - Arabidopsis thaliana  
 N:Alternate names: protein MAA21.70  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 09-Jul-2004  
 C:Accession: T49185  
 R:Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Rudd, S.; Lemcke, K.  
 submitted to the Protein Sequence Database, April 2000  
 A:Reference number: Z25018  
 A:Accession: T49185  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-504 <RIE>  
 A:Cross-references: UNIPROT:Q9LY71; EMBL:AL163818; GSPDB:GNO0061; ATSP:MAA21.70  
 A:Experimental source: cultivar Columbia; BAC clone MAA21  
 C:Genetics:  
 A:Gene: ATSP.MAA21.70  
 A:Map position: 3  
 A:Introns: 196/1; 238/3; 310/3; 398/2  
 C:Superfamily: poppy reticuline oxidase

Query Match 54.4%; Score 80; DB 2; Length 504;  
 Best Local Similarity 46.2%; Pred. No. 0.00036; Mismatches 9; Indels 0; Gaps 0;  
 Matches 12; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

QY 1 SASGLALYPTNRKNDNRMSAMPE 26  
 ::|||:|||||:|||||:|  
 Db 384 TSNGPVIVPVNKSNDNTSAVTP 409

RESULT 4  
 B96785  
 hypothetical protein F1B16.2 [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
 C:Accession: B96785  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
 Jensen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712  
 A:Accession: B96785  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-512 <STO>  
 A:Cross-references: UNIPROT:Q9FWT3; GB:AE005173; NID:g10120443; PIDN:AAG13068.1; GSPDB:G  
 C:Genetics:  
 A:Gene: F1B16.2  
 A:Map position: 1

Query Match 52.4%; Score 77; DB 2; Length 512;  
 Best Local Similarity 54.2%; Pred. No. 0.001; Mismatches 6; Indels 0; Gaps 0;  
 Matches 13; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 SGLALYPTNRKNDNRMSAMPE 26  
 |||:|||||:|||||:|  
 Db 381 SGPILYPMKNDNRSSAVTPD 404

RESULT 5  
 T01500  
 cytochrome oxidase 1 - maize  
 C:Species: Zea mays (maize)  
 C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004  
 C:Accession: T01500  
 R:Morris, R.O.; Laskey, J.G.  
 submitted to the EMBL Data Library, January 1998  
 A:Description: A glycosylated cytochrome oxidase from maize.  
 A:Reference number: Z14336  
 A:Accession: T01500  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-534 <MOR>  
 A:Cross-references: UNIPROT:Q9TON8; EMBL:AF044603; NID:g3341977; PIDN:AAC27500.1; PID:G  
 C:Genetics:  
 A:Gene: cck1  
 A:Introns: 206/1; 437/2

Query Match 42.9%; Score 63; DB 2; Length 534;  
 Best Local Similarity 44.0%; Pred. No. 0.15; Mismatches 5; Indels 0; Gaps 0;  
 Matches 11; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

QY 4 GLALYPTNRKNDNRMSAMPEID 28  
 |||:|||||:|||||:|  
 Db 426 GPLIVYPLNKSNDNRMSAATPSED 450

RESULT 6  
 T51929  
 cytochrome oxidase [imported] - maize  
 C:Species: Zea mays (maize)  
 C:Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 09-Jul-2004  
 C:Accession: T51929  
 R:Houba-Herlin, N.; Pethe, C.; D'Alayer, J.; Laloue, M.  
 Plant J. 17, 615-626, 1999  
 A:Title: Cytochrome oxidase from Zea mays : purification, cDNA cloning and expression in  
 A:Reference number: Z25869  
 A:Accession: T51929  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-534 <HOU>  
 A:Cross-references: UNIPROT:Q9TON8; EMBL:Y18377; PIDN:CAA77151.1  
 A:Experimental source: cultivar nobilis  
 C:Genetics:  
 A:Gene: cko

Query Match 42.9%; Score 63; DB 2; Length 534;  
 Best Local Similarity 44.0%; Pred. No. 0.15; Mismatches 5; Indels 0; Gaps 0;  
 Matches 11; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

QY 4 GLALYPTNRKNDNRMSAMPEID 28  
 |||:|||||:|||||:|  
 Db 426 GPLIVYPLNKSNDNRMSAATPSED 450

RESULT 7  
 T00807  
 probable cytochrome oxidase [imported] - Arabidopsis thaliana  
 N:Alternate names: hypothetical protein T32G6.3  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004  
 C:Accession: T00807; F84842  
 R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul  
 submitted to the EMBL Data Library, November 1997  
 A:Description: Arabidopsis thaliana chromosome II BAC T32G6 genomic sequence.  
 A:Reference number: Z14163  
 A:Accession: T00807  
 A>Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-575 <ROU>  
 A:Cross-references: UNIPROT:O22213; EMBL:AC002510; NID:g2618683; PID:g2618686  
 A:Experimental source: cultivar Columbia

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 22, 2005, 22:34:20 ; Search time 39 Seconds  
(without alignments)  
69.079 Million cell updates/sec

Title: US-10-014-101B-32

Perfect score: 147

Sequence: 1 SASGLALYPTNRKWNMSAMTPEID 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 791.\*

1: Pirl.\*

2: Pirl.\*

3: Pirl.\*

4: Pirl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	130	88.4	515	2 E84577	probable cytokinin
2	108	73.5	524	2 T09937	cytokinin oxidase
3	80	54.4	504	2 T49185	cytokinin oxidase
4	77	52.4	512	2 B96785	hypothetical prote
5	63	42.9	534	2 T01500	cytokinin oxidase
6	63	42.9	534	2 T51829	cytokinin oxidase
7	58	39.5	575	2 T00807	probable cytokinin
8	54	36.7	204	2 B98131	hypothetical prote
9	54	36.7	204	2 AG3156	conserved hypothet
10	53.5	36.4	668	2 S59786	hypothetical prote
11	53	36.1	688	2 D84248	hypothetical prote
12	51.5	35.0	386	2 AG3112	glycerol kinase (E
13	49.5	33.7	523	2 T04332	glutamate-cysteine
14	49	33.3	80	2 H69824	hypothetical prote
15	49	33.3	280	2 B82099	phosphatidate cyti
16	49	33.3	466	1 IQEBV	replication initia
17	48	32.7	240	2 S75162	hypothetical prote
18	48	32.7	246	2 F97076	hypothetical prote
19	48	32.7	326	1 JQ1443	glycoprotein VP7 p
20	48	32.7	1070	2 C75506	hypothetical prote
21	47.5	32.3	227	2 F90622	ATP synthase F0 ch
22	47	32.0	156	2 AB0079	hypothetical prote
23	47	32.0	784	2 T45697	hypothetical prote
24	47	32.0	1051	2 S59791	probable membrane
25	47	32.0	1324	2 T01508	mismatch repair en
26	46.5	31.6	224	2 F70798	hypothetical prote
27	46.5	31.6	255	2 A87619	hypothetical prote
28	46	31.3	338	2 S75196	hypothetical prote
29	46	31.3	494	2 B83351	conserved hypothet

30	46	31.3	890	2 T30103	hypothetical prote
31	45.5	31.0	227	2 T11027	H+-transporting tw
32	45.5	31.0	227	2 T11185	H+-transporting tw
33	45.5	31.0	227	2 T10991	H+-transporting tw
34	45.5	31.0	287	2 AG2414	hypothetical prote
35	45.5	31.0	568	2 T05218	hypothetical prote
36	45	30.6	350	2 D71348	flagellar filament
37	44.5	30.3	209	2 S31071	rpcF protein - syn
38	44.5	30.3	499	2 AD3035	glycerol kinase gl
39	44.5	30.3	507	2 G98250	probable carbohydr
40	44.5	30.3	872	2 E71852	valine-tRNA ligase
41	44.5	30.3	874	2 A64664	valine-tRNA ligase
42	44.5	30.3	1075	2 AF2355	hypothetical prote
43	44	29.9	116	2 F82841	hypothetical prote
44	44	29.9	152	2 A41222	ubiquitin-protein
45	44	29.9	152	2 B41222	ubiquitin-protein

ALIGNMENTS

RESULT 1

E84577

probable cytokinin oxidase [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cross)

C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 19-Apr-2002

C;Accession: E84577

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
euss, D.; Niernman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.W.; Venter, C.

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: E84577

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-515 <STO>

A;Cross-references: GB:AE002093; NID:g4191780; PIDN:AA010149.1; GSPDB:GN00139

C;Genetics:

A;Gene: At2g19500

A;Map position: 2

C;Superfamily: poppy reticuline oxidase

Query Match 88.4%; Score 130; DB 2; Length 515;

Best Local Similarity 66.7%; Pred. No. 8.9e-12;

Matches 28; Conservative 0; Mismatches 0; Indels 14; Gaps 1;

QY 1 SASGLALYPTNRN-----KWNMSAMTPEID 28

DB 391 SASGLALYPTNRNKYILLFIHYLQEPKWNMSAMTPEID 432

RESULT 2

T09937

cytokinin oxidase homolog T16L4.250 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cross)

C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004

C;Accession: T09937

R;Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.

submitted to the Protein Sequence Database, June 1999

A;Reference number: Z16897

A;Accession: T09937

A;Molecule type: DNA

A;Residues: 1-524 <BEV>

A;Cross-references: UNIPROT:O9PUJ2; EMBL:AL079344; GSPDB:GN00062; ATSP:T16L4.250

A;Experimental source: cultivar Columbia; BAC clone T16L4

C;Genetics:

A;Gene: ATSP:T16L4.250

A;Map position: 4

A;Introns: 210/1; 252/3; 338/3; 428/2

Query Match 73.5%; Score 108; DB 2; Length 524;

Best Local Similarity 67.9%; Pred. No. 2e-08;

**This Page Blank (uspto)**

Wed Feb 23 08:35:34 2005

```
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 146464
; LENGTH: 558
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_47086C.1.pep
US-10-437-963-146464

Query Match          59.2%; Score 87; DB 16; Length 558;
Best Local Similarity 59.3%; Pred. No. 6.6e-05;
Matches 16; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy      2 ASGLALYPTNRKNDNRMSAMIPED 28
Db      440 AMGFVLIYPMNRKNKWDNSNSAVITDDD 466

RESULT 15
US-10-425-114-40061
; Sequence 40061, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 40061
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700978152_FLI.pep
US-10-425-114-40061

Query Match          57.8%; Score 85; DB 15; Length 292;
Best Local Similarity 57.7%; Pred. No. 6.7e-05;
Matches 15; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy      3 SGLALYPTNRKNDNRMSAMIPED 28
Db      172 SGPIIYPMNRKNKWDHRSSVVTPEED 197
```

Search completed: February 22, 2005, 22:56:47  
Job time : 131 secs

```
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 39759
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700731265_FLI.pep
US-10-425-114-39759

Query Match          61.9%; Score 91; DB 15; Length 290;
Best Local Similarity 53.6%; Pred. No. 7.5e-06;
Matches 15; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 SASGLALLYPTNRNKNWDRMSAMPEID 28
Db 170 TSNGLVLIYPVNRKSKWDRNTSVIPEED 197

RESULT 11
US-10-424-599-144567
; Sequence 144567, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5323)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 144567
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_101557C.1.pep
US-10-424-599-144567

Query Match          61.9%; Score 91; DB 15; Length 324;
Best Local Similarity 53.6%; Pred. No. 8.5e-06;
Matches 15; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 SASGLALLYPTNRNKNWDRMSAMPEID 28
Db 204 TSNGLVLIYPVNRKSKWDRNTSVIPEED 231

RESULT 12
US-10-425-114-42126
; Sequence 42126, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
```

```
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 42126
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700985454_FLI.pep
US-10-425-114-42126

Query Match          61.9%; Score 91; DB 15; Length 324;
Best Local Similarity 53.6%; Pred. No. 8.5e-06;
Matches 15; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 SASGLALLYPTNRNKNWDRMSAMPEID 28
Db 204 TSNGLVLIYPVNRKSKWDRNTSVIPEED 231

RESULT 13
US-10-437-963-148492
; Sequence 148492, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 148492
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_48918C.1.pep
US-10-437-963-148492

Query Match          59.9%; Score 88; DB 16; Length 525;
Best Local Similarity 61.5%; Pred. No. 4.3e-05;
Matches 16; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 SASGLALLYPTNRNKNWDRMSAMPE 26
Db 404 SNNGLIYPVNRKSKWDRNTSVIPE 429

RESULT 14
US-10-437-963-146464
; Sequence 146464, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
```

```
US-10-311-453-8
; Sequence 8, Application US/10311453
; Publication No. US20040031073A1
; GENERAL INFORMATION:
; APPLICANT: Schmling, Thomas
; APPLICANT: Werner, Tom s
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and
; FILE REFERENCE: 1226-4
; CURRENT APPLICATION NUMBER: US/10/311,453
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: EP 00870132.8
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/258,415
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: EP 01870053.4
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-311-453-8

Query Match      73.5%; Score 108; DB 15; Length 524;
Best Local Similarity 67.9%; Pred. No. 3e-08;
Matches 19; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 SASGLALYPTNRKNDNRMSAMPEID 28
Db 414 STSGVTLFYPTNRKNDNRMSAMPEID 441

RESULT 7
US-10-014-101-6
; Sequence 6, Application US/10014101
; Publication No. US20030074698A1
; GENERAL INFORMATION:
; APPLICANT: Schmling, Thomas
; APPLICANT: Werner, Tom s
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and
; FILE REFERENCE: 1195-2
; CURRENT APPLICATION NUMBER: US/10/014,101
; CURRENT FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: PCT/EP01/06833
; PRIOR FILING DATE: 2001-06-16
; PRIOR APPLICATION NUMBER: EP 00870132.8
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/258,415
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: EP 01870053.4
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-014-101-6

Query Match      65.3%; Score 96; DB 14; Length 523;
Best Local Similarity 69.2%; Pred. No. 2.4e-06;
Matches 18; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 3 SGALLYPTNRKNDNRMSAMPEID 28
Db 411 SGFVLVYPMNRKNDNRMSAAIPEED 436

RESULT 8
US-10-311-453-6
; Sequence 8, Application US/10311453
; Publication No. US20040031073A1
; GENERAL INFORMATION:
; APPLICANT: Schmling, Thomas
; APPLICANT: Werner, Tom s
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and
; FILE REFERENCE: 1226-4
; CURRENT APPLICATION NUMBER: US/10/311,453
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: EP 00870132.8
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/258,415
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: EP 01870053.4
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-311-453-6

Query Match      73.5%; Score 108; DB 15; Length 524;
Best Local Similarity 67.9%; Pred. No. 3e-08;
Matches 19; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 SASGLALYPTNRKNDNRMSAMPEID 28
Db 414 STSGVTLFYPTNRKNDNRMSAMPEID 441

RESULT 7
US-10-014-101-6
; Sequence 6, Application US/10014101
; Publication No. US20030074698A1
; GENERAL INFORMATION:
; APPLICANT: Schmling, Thomas
; APPLICANT: Werner, Tom s
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and
; FILE REFERENCE: 1195-2
; CURRENT APPLICATION NUMBER: US/10/014,101
; CURRENT FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: PCT/EP01/06833
; PRIOR FILING DATE: 2001-06-16
; PRIOR APPLICATION NUMBER: EP 00870132.8
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/258,415
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: EP 01870053.4
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-014-101-6

Query Match      65.3%; Score 96; DB 14; Length 523;
Best Local Similarity 69.2%; Pred. No. 2.4e-06;
Matches 18; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 3 SGALLYPTNRKNDNRMSAMPEID 28
Db 411 SGFVLVYPMNRKNDNRMSAAIPEED 436

RESULT 8
US-10-311-453-6
; Sequence 8, Application US/10311453
; Publication No. US20040031073A1
; GENERAL INFORMATION:
; APPLICANT: Schmling, Thomas
; APPLICANT: Werner, Tom s
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and
; FILE REFERENCE: 1226-4
; CURRENT APPLICATION NUMBER: US/10/311,453
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: EP 00870132.8
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/258,415
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: EP 01870053.4
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-311-453-6

Query Match      65.3%; Score 96; DB 15; Length 523;
Best Local Similarity 69.2%; Pred. No. 2.4e-06;
Matches 18; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 3 SGALLYPTNRKNDNRMSAMPEID 28
Db 411 SGFVLVYPMNRKNDNRMSAAIPEED 436

RESULT 9
US-10-424-599-144566
; Sequence 144566, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 144566
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_101556C.1.pep
US-10-424-599-144566

Query Match      61.9%; Score 91; DB 15; Length 290;
Best Local Similarity 53.6%; Pred. No. 7.5e-06;
Matches 15; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

Qy 1 SASGLALYPTNRKNDNRMSAMPEID 28
Db 170 TSGFVLVYPMNRKNDNRMSAAIPEED 197

RESULT 10
US-10-425-114-39759
; Sequence 39759, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
```

```

RESULT 2
US-10-311-453-32
; Sequence 32, Application US/10311453
; Publication No. US20040031073A1
; GENERAL INFORMATION:
; APPLICANT: Schmulling, Thomas
; APPLICANT: Werner, Tom s
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and
; FILE OF INVENTION: physiology
; FILE REFERENCE: 1226-4
; CURRENT APPLICATION NUMBER: US/10/311,453
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: EP 00870132.8
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/258,415
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: EP 01870053.4
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 32
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-311-453-32

Query Match 100.0%; Score 147; DB 15; Length 28;
Best Local Similarity 100.0%; Pred. No. 8,7e-16;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SASGLALYPTNRNKNWDRMSAMIPED 28
DB 1 SASGLALYPTNRNKNWDRMSAMIPED 28

RESULT 3
US-10-014-101-4
; Sequence 4, Application US/10014101
; Publication No. US20030074698A1
; GENERAL INFORMATION:
; APPLICANT: Schmulling, Thomas
; APPLICANT: Werner, Tom s
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and
; FILE OF INVENTION: physiology
; FILE REFERENCE: 1195-2
; CURRENT APPLICATION NUMBER: US/10/014,101
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: PCT/EP01/06833
; PRIOR FILING DATE: 2001-06-16
; PRIOR APPLICATION NUMBER: EP 00870132.8
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/258,415
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: EP 01870053.4
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-014-101-4

Query Match 100.0%; Score 147; DB 14; Length 501;
Best Local Similarity 100.0%; Pred. No. 2,1e-14;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SASGLALYPTNRNKNWDRMSAMIPED 28
DB 391 SASGLALYPTNRNKNWDRMSAMIPED 418

RESULT 4
US-10-311-453-4
; Sequence 4, Application US/10311453
; Publication No. US20040031073A1
; GENERAL INFORMATION:
; APPLICANT: Schmulling, Thomas
; APPLICANT: Werner, Tom s
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and
; FILE OF INVENTION: physiology
; FILE REFERENCE: 1226-4
; CURRENT APPLICATION NUMBER: US/10/311,453
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: EP 00870132.8
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/258,415
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: EP 01870053.4
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-311-453-4

Query Match 100.0%; Score 147; DB 15; Length 501;
Best Local Similarity 100.0%; Pred. No. 2,1e-14;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SASGLALYPTNRNKNWDRMSAMIPED 28
DB 391 SASGLALYPTNRNKNWDRMSAMIPED 418

RESULT 5
US-10-014-101-8
; Sequence 8, Application US/10014101
; Publication No. US20030074698A1
; GENERAL INFORMATION:
; APPLICANT: Schmulling, Thomas
; APPLICANT: Werner, Tom s
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and
; FILE OF INVENTION: physiology
; FILE REFERENCE: 1195-2
; CURRENT APPLICATION NUMBER: US/10/014,101
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: PCT/EP01/06833
; PRIOR FILING DATE: 2001-06-16
; PRIOR APPLICATION NUMBER: EP 00870132.8
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/258,415
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: EP 01870053.4
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-014-101-8

Query Match 73.5%; Score 108; DB 14; Length 524;
Best Local Similarity 67.9%; Pred. No. 3e-08;
Matches 19; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 SASGLALYPTNRNKNWDRMSAMIPED 28
DB 414 STSGVTLFYPTNRNKNWDRMSAMIPED 441

RESULT 6
US-10-014-101-8
; Sequence 8, Application US/10014101
; Publication No. US20030074698A1
; GENERAL INFORMATION:
; APPLICANT: Schmulling, Thomas
; APPLICANT: Werner, Tom s
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and
; FILE OF INVENTION: physiology
; FILE REFERENCE: 1195-2
; CURRENT APPLICATION NUMBER: US/10/014,101
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: PCT/EP01/06833
; PRIOR FILING DATE: 2001-06-16
; PRIOR APPLICATION NUMBER: EP 00870132.8
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/258,415
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: EP 01870053.4
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-014-101-8

Query Match 73.5%; Score 108; DB 14; Length 524;
Best Local Similarity 67.9%; Pred. No. 3e-08;
Matches 19; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 SASGLALYPTNRNKNWDRMSAMIPED 28
DB 414 STSGVTLFYPTNRNKNWDRMSAMIPED 441

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 22, 2005, 22:44:21 ; Search time 130 Seconds  
(without alignments)  
70.483 Million cell updates/sec

Title: US-10-014-101B-32

Perfect score: 147

Sequence: 1 SASGLALYPTRNRKNWNRMSAMPEID 28

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1380268 seqs, 327241040 residues

Total number of hits satisfying chosen parameters: 1380268

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

```
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	147	100.0	28	14 US-10-014-101-32	Sequence 32, Appl
2	147	100.0	28	15 US-10-311-453-32	Sequence 32, Appl
3	147	100.0	501	14 US-10-014-101-4	Sequence 4, Appl
4	147	100.0	501	15 US-10-311-453-4	Sequence 4, Appl
5	108	73.5	524	14 US-10-014-101-8	Sequence 8, Appl
6	108	73.5	524	15 US-10-311-453-8	Sequence 8, Appl
7	96	65.3	523	14 US-10-014-101-6	Sequence 6, Appl
8	96	65.3	523	15 US-10-311-453-6	Sequence 6, Appl
9	91	61.9	230	15 US-10-424-599-144566	Sequence 144566, A
10	91	61.9	230	15 US-10-425-114-39759	Sequence 39759, A
11	91	61.9	324	15 US-10-424-599-144567	Sequence 144567, A
12	91	61.9	324	15 US-10-425-114-42126	Sequence 42126, A
13	88	59.9	525	16 US-10-437-963-148492	Sequence 148492, A

14	87	59.2	558	16	US-10-437-963-146464	Sequence 146464, A
15	85	57.8	292	15	US-10-425-114-40061	Sequence 40061, A
16	85	57.8	299	15	US-10-424-599-162425	Sequence 162425, A
17	85	57.8	550	16	US-10-437-963-186310	Sequence 186310, A
18	83	56.5	173	15	US-10-425-114-52875	Sequence 52875, A
19	83	56.5	521	16	US-10-437-963-173054	Sequence 173054, A
20	82	55.8	143	15	US-10-424-599-205473	Sequence 205473, A
21	81	55.1	534	16	US-10-437-963-140653	Sequence 140653, A
22	80	54.4	504	14	US-10-014-101-12	Sequence 12, Appl
23	80	54.4	504	15	US-10-311-453-12	Sequence 12, Appl
24	77	52.4	536	14	US-10-014-101-10	Sequence 10, Appl
25	77	52.4	536	15	US-10-311-453-10	Sequence 10, Appl
26	77	52.4	539	14	US-10-014-101-35	Sequence 35, Appl
27	77	52.4	539	15	US-10-311-453-35	Sequence 35, Appl
28	75	51.0	214	15	US-10-424-599-200943	Sequence 200943, A
29	75	51.0	527	16	US-10-437-963-136938	Sequence 136938, A
30	74	50.3	532	16	US-10-437-963-183757	Sequence 183757, A
31	68	46.3	517	16	US-10-437-963-177713	Sequence 177713, A
32	66	44.9	392	15	US-10-425-114-39823	Sequence 39823, A
33	66	44.9	513	15	US-10-424-599-220961	Sequence 220961, A
34	64	43.5	229	15	US-10-425-114-39997	Sequence 39997, A
35	64	43.5	230	15	US-10-424-599-255266	Sequence 255266, A
36	63	42.9	611	9	US-09-829-549A-48	Sequence 48, Appl
37	59	40.1	164	16	US-10-437-963-173076	Sequence 173076, A
38	58	39.5	99	15	US-10-424-599-204492	Sequence 204492, A
39	58	39.5	99	15	US-10-425-114-56907	Sequence 56907, A
40	58	39.5	575	14	US-10-014-101-2	Sequence 2, Appl
41	58	39.5	575	15	US-10-311-453-2	Sequence 2, Appl
42	57	38.8	532	16	US-10-437-963-191047	Sequence 191047, A
43	51	34.7	395	16	US-10-437-963-204083	Sequence 204083, A
44	51	34.7	572	16	US-10-437-963-189286	Sequence 189286, A
45	50	34.0	248	16	US-10-437-963-185716	Sequence 185716, A

#### ALIGNMENTS

##### RESULT 1

```
US-10-014-101-32
; Sequence 32, Application US/10014101
; Publication No. US20030074698A1
; GENERAL INFORMATION:
; APPLICANT: Schmulling, Thomas
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and
; TITLE OF INVENTION: physiology
; FILE REFERENCE: 1195-2
; CURRENT APPLICATION NUMBER: US/10/014,101
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: PCT/EP01/06833
; PRIOR FILING DATE: 2001-06-16
; PRIOR APPLICATION NUMBER: EP 00870132.8
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/258,415
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: EP 01870053.4
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-014-101-32
```

Query Match 100.0% Score 147; DB 14; Length 28;  
Best Local Similarity 100.0%; Pred. No. 8.7e-16;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SASGLALYPTRNRKNWNRMSAMPEID 28

DB 1 SASGLALYPTRNRKNWNRMSAMPEID 28

This Page Blank (115010)

```
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-9994

Query Match      30.6%; Score 45; DB 4; Length 286;
Best Local Similarity 36.4%; Pred. No. 40;
Matches 8; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY      4 GLALLYPTNRNKNWDRMSAMIP 25
      |||||::|||::|
Db      225 GAALVTFGLGHRVWVRLPLLP 246

RESULT 15
US-09-270-767-46433
; Sequence 46433, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46433
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-46433

Query Match      30.6%; Score 45; DB 4; Length 314;
Best Local Similarity 45.5%; Pred. No. 45;
Matches 10; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY      2 ASGLALLYPTNRNKNWDRMSAM 23
      |||||::|||::|
Db      288 AFNLHLIYRANGNKKARQYLM 309
```

Search completed: February 22, 2005, 22:45:48  
Job time : 43 secs

Db 427 ARLYNTNEANANARLSAMLPYV 448

## RESULT 12

US-07-973-320-2  
; Sequence 2, Application US/07973320  
; Patent No. 5286486  
; GENERAL INFORMATION:  
; APPLICANT: Payne, Jewel M.  
; APPLICANT: Fu, Jenny M.  
; TITLE OF INVENTION: No. 5286486el Bacillus thuringiensis Gene  
; TITLE OF INVENTION: Encoding a Coleopteran-Active Toxin  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David R. Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: USA  
; ZIP: 32606

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/973,320  
; FILING DATE: 19921106  
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/788,638  
; FILING DATE: 6-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Saliwanchik, David R.  
; REGISTRATION NUMBER: 31,794  
; REFERENCE/DOCKET NUMBER: MA68.C1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 904-375-8100

TELEFAX: 904-372-5800  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:

LENGTH: 1138 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single

TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES

ANTI-SENSE: NO  
; ORIGINAL SOURCE:

ORGANISM: Bacillus thuringiensis  
; STRAIN: dakota

INDIVIDUAL ISOLATE: HD511  
; IMMEDIATE SOURCE:

LIBRARY: Lamdagem (TM)-11 library of J.M. Fu  
; CLONE: 511

US-07-973-320-2

Query Match 31.3%; Score 46; DB 1; Length 1138;  
Best Local Similarity 53.3%; Pred. No. 1.4e+02;  
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 7 LLYPTNRNKNWDRMS 21

||:|:|:|:|:|:|:  
Db 91 LLWPHNKNIWDEFMT 105

## RESULT 13

US-07-973-320-4  
; Sequence 4, Application US/07973320  
; Patent No. 5286486  
; GENERAL INFORMATION:  
; APPLICANT: Payne, Jewel M.  
; APPLICANT: Fu, Jenny M.

; TITLE OF INVENTION: No. 5286486el Bacillus thuringiensis Gene  
; TITLE OF INVENTION: Encoding a Coleopteran-Active Toxin  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David R. Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: USA  
; ZIP: 32606

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/973,320  
; FILING DATE: 19921106  
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/788,638  
; FILING DATE: 6-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Saliwanchik, David R.

REGISTRATION NUMBER: 31,794  
; REFERENCE/DOCKET NUMBER: MA68.C1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 904-375-8100

TELEFAX: 904-372-5800  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:

LENGTH: 1138 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single

TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES

ANTI-SENSE: NO  
; ORIGINAL SOURCE:

ORGANISM: Bacillus thuringiensis  
; STRAIN: kumamotoensis  
; INDIVIDUAL ISOLATE: HD867

IMMEDIATE SOURCE:  
; LIBRARY: Lamdagem (TM)-11 library of J.M. Fu  
; CLONE: 867

US-07-973-320-4

Query Match 31.3%; Score 46; DB 1; Length 1138;  
Best Local Similarity 53.3%; Pred. No. 1.4e+02;  
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 7 LLYPTNRNKNWDRMS 21

||:|:|:|:|:|:|:  
Db 91 LLWPHNKNIWDEFMT 105

## RESULT 14

US-09-902-540-9994  
; Sequence 9994, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.

; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; CURRENT FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 9994

## RESULT 15

US-09-902-540-9994  
; Sequence 9994, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.

; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; CURRENT FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 9994

## RESULT 16

US-09-902-540-9994  
; Sequence 9994, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.

; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; CURRENT FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 9994

```

RESULT 7
US-09-911-882-14
; Sequence 14, Application US/09911882
; Patent No. 6465198
; GENERAL INFORMATION:
; APPLICANT: Koltin, Yigal
; APPLICANT: Gavrias, Victoria
; TITLE OF INVENTION: ESSENTIAL FUNGAL GENES AND THEIR USE
; FILE REFERENCE: 06286-062004
; CURRENT APPLICATION NUMBER: US/09/911,882
; CURRENT FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: US 08/965,762
; PRIOR FILING DATE: 1997-11-07
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 1051
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-911-882-14

Query Match 32.0%; Score 47; DB 4; Length 1051;
Best Local Similarity 53.8%; Pred. No. 90;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 16 WDNRMASAMPEID 28
||| :|||:
Db 796 WDDLQSSVPEVD 808

RESULT 8
US-09-911-888-14
; Sequence 14, Application US/09911888
; Patent No. 6514715
; GENERAL INFORMATION:
; APPLICANT: Koltin, Yigal
; APPLICANT: Gavrias, Victoria
; TITLE OF INVENTION: ESSENTIAL FUNGAL GENES AND THEIR USE
; FILE REFERENCE: 06286-062002
; CURRENT APPLICATION NUMBER: US/09/911,888
; CURRENT FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: US 08/965,762
; PRIOR FILING DATE: 1997-11-07
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 1051
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-911-888-14

Query Match 32.0%; Score 47; DB 4; Length 1051;
Best Local Similarity 53.8%; Pred. No. 90;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 16 WDNRMASAMPEID 28
||| :|||:
Db 796 WDDLQSSVPEVD 808

RESULT 9
US-09-270-767-60682
; Sequence 60682, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 60682
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-60682

Query Match 31.3%; Score 46; DB 4; Length 280;
Best Local Similarity 43.8%; Pred. No. 27;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 10 PTRNKNWNRMSAMIP 25
||| :|||:
Db 245 PIRDLFDSALEVMVP 260

RESULT 10
US-09-248-796A-15064
; Sequence 15064, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 15064
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-15064

Query Match 31.3%; Score 46; DB 4; Length 466;
Best Local Similarity 32.1%; Pred. No. 50;
Matches 9; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY 1 SASGLALYPTNRNKNWNRMSAMIDEI 28
||| :|||:
Db 407 SASALLQOALTKQWLNRSLILQSV 434

RESULT 11
US-09-252-991A-25323
; Sequence 25323, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25323
; LENGTH: 551
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25323

Query Match 31.3%; Score 46; DB 4; Length 551;
Best Local Similarity 45.5%; Pred. No. 61;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 6 ALLYPTNRNKNWNRMSAMIDEI 27
||| :|||:
; SEQ ID NO 60682
```

```
Best Local Similarity 44.0%; Pred. No. 0.11;
Matches 11; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 4 GLALLYPTNRNKKWDRMSAMPEID 28
DB 426 GPLIVYPLNKSMDGMSAATPSD 450

RESULT 3
US-08-728-603-19
; Sequence 19, Application US/08728603
; Patent No. 6093806
; GENERAL INFORMATION:
; APPLICANT: Cesarman, Ethel
; APPLICANT: Knowles, Daniel M.
; TITLE OF INVENTION: PROTEINS OF KAPOSI'S SARCOMA ASSOCIATED
; TITLE OF INVENTION: HERPESVIRUS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/728,603
; FILING DATE: 10-OCT-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BRAMAN, SUSAN J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 19603/720
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 716-263-1636
; TELEFAX: 716-263-1600
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-728-603-19

Query Match 33.3%; Score 49; DB 3; Length 257;
Best Local Similarity 37.0%; Pred. No. 8.1;
Matches 10; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

QY 1 SASGLALLYPTNRNKKWDRMSAMPEI 27
DB 198 SAAGCALLVPANVIPDTHSGGVVQOL 224

RESULT 4
US-09-543-681A-7393
; Sequence 7393, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
```

```
; SEQ ID NO 7393
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7393

Query Match 33.3%; Score 49; DB 4; Length 470;
Best Local Similarity 39.1%; Pred. No. 17;
Matches 9; Conservative 5; Mismatches 5; Indels 4; Gaps 1;

QY 10 PTN-----RNKWDNRMSAMPEID 28
DB 108 PTNSQVPSPSWNDNQPSQLPELN 130

RESULT 5
US-08-965-762-14
; Sequence 14, Application US/08965762
; Patent No. 6280963
; GENERAL INFORMATION:
; APPLICANT: Koltin, Yigal
; APPLICANT: Gavrias, Victoria
; TITLE OF INVENTION: ESSENTIAL FUNGAL GENES AND THEIR USE
; FILE REFERENCE: 07334/062001
; CURRENT APPLICATION NUMBER: US/08/965,762
; CURRENT FILING DATE: 1997-11-07
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 1051
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-08-965-762-14

Query Match 32.0%; Score 47; DB 3; Length 1051;
Best Local Similarity 53.8%; Pred. No. 90;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 16 WDRNRSAMPEID 28
DB 796 WDDLQSSVPEVD 808

RESULT 6
US-09-911-927-14
; Sequence 14, Application US/09911927
; Patent No. 6461826
; GENERAL INFORMATION:
; APPLICANT: Koltin, Yigal
; APPLICANT: Gavrias, Victoria
; TITLE OF INVENTION: ESSENTIAL FUNGAL GENES AND THEIR USE
; FILE REFERENCE: 06286-062003
; CURRENT APPLICATION NUMBER: US/09/911,927
; CURRENT FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: US 08/965,762
; PRIOR FILING DATE: 1997-11-07
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 1051
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-911-927-14

Query Match 32.0%; Score 47; DB 4; Length 1051;
Best Local Similarity 53.8%; Pred. No. 90;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 16 WDRNRSAMPEID 28
DB 796 WDDLQSSVPEVD 808
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 22, 2005, 22:35:25 ; Search time 42 Seconds  
(without alignments)  
49.766 Million cell updates/sec

Title: US-10-014-101B-32

Perfect score: 147  
Sequence: 1 SASGLALYPTNRKNRMSAMPEID 28

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS COMB pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	63	42.9	534	3 US-09-124-541-1	Sequence 1, Appli
2	63	42.9	534	4 US-09-663-326-1	Sequence 1, Appli
3	49	33.3	257	3 US-08-728-603-19	Sequence 19, Appl
4	49	33.3	470	4 US-09-543-681A-7393	Sequence 7393, Ap
5	47	32.0	1051	3 US-08-965-762-14	Sequence 14, Appl
6	47	32.0	1051	4 US-09-911-927-14	Sequence 14, Appl
7	47	32.0	1051	4 US-09-911-882-14	Sequence 14, Appl
8	47	32.0	1051	4 US-09-911-888-14	Sequence 14, Appl
9	46	31.3	280	4 US-09-270-767-60682	Sequence 60682, A
10	46	31.3	466	4 US-09-248-796A-15064	Sequence 15064, A
11	46	31.3	551	4 US-09-252-991A-25323	Sequence 25323, A
12	46	31.3	1138	1 US-07-973-320-2	Sequence 2, Appli
13	46	31.3	1138	1 US-07-973-320-4	Sequence 4, Appli
14	45	30.6	286	4 US-09-902-540-9994	Sequence 9994, Ap
15	45	30.6	314	4 US-09-270-767-46433	Sequence 46433, A
16	45	30.6	938	4 US-09-602-787A-172	Sequence 172, App
17	44.5	30.3	795	4 US-09-248-796A-16677	Sequence 16677, A
18	44.5	30.3	872	1 US-08-451-715A-8	Sequence 8, Appli
19	44	29.9	130	4 US-09-621-976-4511	Sequence 4511, Ap
20	44	29.9	147	4 US-09-602-787A-252	Sequence 252, App
21	44	29.9	152	1 US-08-318-947A-6	Sequence 6, Appli
22	44	29.9	152	1 US-08-318-947A-7	Sequence 7, Appli
23	44	29.9	152	1 US-08-318-947A-8	Sequence 8, Appli
24	44	29.9	152	2 US-08-795-303-6	Sequence 6, Appli
25	44	29.9	152	2 US-08-795-303-7	Sequence 7, Appli
26	44	29.9	152	2 US-08-795-303-8	Sequence 8, Appli
27	44	29.9	152	2 US-08-247-904B-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1

US-09-124-541-1  
; Sequence 1, Application US/09124541A  
; Patent No. 6229066  
; GENERAL INFORMATION:  
; APPLICANT: Morris Ph.D., Roy O.  
; TITLE OF INVENTION: A CYTOKININ OXIDASE  
; FILE REFERENCE: UMO1490  
; CURRENT APPLICATION NUMBER: US/09/124,541A  
; CURRENT FILING DATE: 1998-07-29  
; EARLIER APPLICATION NUMBER: 60/054,268  
; EARLIER FILING DATE: 1997-07-30  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 534  
; TYPE: PRT  
; ORGANISM: Zea mays  
US-09-124-541-1

Query Match 42.9%; Score 63; DB 3; Length 534;  
Best Local Similarity 44.0%; Pred. No. 0.11;  
Matches 11; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 4 GLALYPTNRKNRMSAMPEID 28  
DB 426 GPLIVYPLNKGWDGMSAATPSED 450

RESULT 2

US-09-663-326-1  
; Sequence 1, Application US/09663326  
; Patent No. 6617497  
; GENERAL INFORMATION:  
; APPLICANT: Morris Ph.D., Roy O.  
; TITLE OF INVENTION: A CYTOKININ OXIDASE  
; FILE REFERENCE: UMO1490  
; CURRENT APPLICATION NUMBER: US/09/663,326  
; CURRENT FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: 60/054,268  
; PRIOR FILING DATE: 1997-07-30  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 534  
; TYPE: PRT  
; ORGANISM: Zea mays  
US-09-663-326-1

Query Match 42.9%; Score 63; DB 4; Length 534;

Sequence 17, Appl  
Sequence 1000, Ap  
Sequence 1155, Ap  
Sequence 8138, Ap  
Sequence 11571, A  
Sequence 5105, Ap  
Sequence 1, Appli  
Sequence 4700, Ap  
Sequence 26352, A  
Sequence 12, Appl  
Sequence 13, Appl  
Sequence 31564, A  
Sequence 42050, A  
Sequence 4707, Ap  
Sequence 14154, A  
Sequence 2, Appli  
Sequence 21456, A  
Sequence 18601, A

PD 20-DEC-2001.  
 XX  
 PF 18-JUN-2001; 2001WO-EF006833.  
 XX  
 PR 16-JUN-2000; 2000EP-00870132.  
 PR 27-DEC-2000; 2000US-0258415P.  
 PR 16-MAR-2001; 2001EP-00870053.  
 XX  
 XX (SCHM/) SCHMULLING T.  
 PA (WERN/) WERNER T.  
 XX  
 PI Schmullling T, Werner T;  
 XX  
 DR WPI; 2002-130736/17.  
 DR N-PSDB; ABK28609.  
 XX  
 PT Polynucleotide encoding novel plant protein having cytokinin oxidase  
 PT activity and the protein useful for stimulating root growth, enhancing  
 PT the formation of lateral or adventitious roots, altering root geotropism.  
 XX  
 PS Claim 2; Page 135-137; 154pp; English.  
 XX  
 CC The invention relates to an isolated polynucleotide (I) encoding a novel  
 CC plant protein (II) having cytokinin oxidase activity. (I) is useful for  
 CC production of transgenic plants, plant cells or tissues; for production  
 CC of altered plants, plant cell or tissues; and for effecting the  
 CC expression of (II) where (I) is operably linked to one or more control  
 CC sequences. The methods further comprises regenerating a plant from the  
 CC plant cell. (I) and (II) are useful for stimulating root growth;  
 CC enhancing the formation of lateral or adventitious roots; altering root  
 CC geotropism, leading to an increase in yield; and for screening growth  
 CC promoting chemical of herbicides. (I) is useful for increasing the size  
 CC of the root meristem; increasing root size; increasing the size of the  
 CC shoot meristem; delaying leaf senescence and altering leaf senescence;  
 CC increasing leaf thickness; reducing or increasing the vessel size;  
 CC inducing parthenocarp; improving standability of the seedlings;  
 CC increasing branching and for improving lodging resistance. Antibody (III)  
 CC to (II) is useful for identifying and obtaining proteins interacting with  
 CC (II) comprising a screening assay, preferably a two-hybrid screening  
 CC assay. AAU81967-AAU81974 represent Arabidopsis thaliana cytokinin oxidase  
 CC amino acid sequences  
 XX  
 SQ Sequence 524 AA;

Query Match 73.5%; Score 108; DB 5; Length 524;  
 Best Local Similarity 67.9%; Pred. NO. 1.1e-08;  
 Matches 19; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 SASGLALLYPTNRNKNWNSAMPEID 28  
 Db 414 STSGVTLFYPTNRNKNWNSMTPTD 441

Search completed: February 22, 2005, 22:41:12  
 Job time : 168 secs



```
OY      1 SASGALLYPTRNRKWNKMSAMPEID 28
Db      290 STSGVTLFYPTNRKWNKMSMTWPD 317
      |||:|||||:|||||:|
RESULT 13
AAG31316
ID  AAG31316 standard; protein; 524 AA.
XX      AAG31316;
XX      17-OCT-2000 (first entry)
XX      Arabidopsis thaliana protein fragment SEQ ID NO: 37589.
DE      Arabidopsis thaliana protein fragment SEQ ID NO: 37589.
XX      Protein identification; signal transduction pathway; metabolic pathway;
KW      hybridisation assay; genetic mapping; gene expression control; promoter;
KW      termination sequence.
XX      Arabidopsis thaliana.
OS      Arabidopsis thaliana.
XX      EP1033405-A2.
PN      06-SEP-2000.
PD      25-FEB-2000; 2000EP-00301439.
PF      25-FEB-1999; 99US-0121825P.
XX      05-MAR-1999; 99US-0123180P.
PR      09-MAR-1999; 99US-0123548P.
PR      23-MAR-1999; 99US-0125788P.
PR      25-MAR-1999; 99US-0126264P.
PR      29-MAR-1999; 99US-0126785P.
PR      01-APR-1999; 99US-0127462P.
PR      06-APR-1999; 99US-0128234P.
PR      08-APR-1999; 99US-0128714P.
PR      16-APR-1999; 99US-0129845P.
PR      19-APR-1999; 99US-0130077P.
PR      21-APR-1999; 99US-0130449P.
PR      23-APR-1999; 99US-0130510P.
PR      28-APR-1999; 99US-0130891P.
PR      30-APR-1999; 99US-0131449P.
PR      30-APR-1999; 99US-0132048P.
PR      30-APR-1999; 99US-0132407P.
PR      04-MAY-1999; 99US-0132484P.
PR      05-MAY-1999; 99US-0132485P.
PR      06-MAY-1999; 99US-0132486P.
PR      06-MAY-1999; 99US-0132487P.
PR      07-MAY-1999; 99US-0132863P.
PR      11-MAY-1999; 99US-0134256P.
PR      14-MAY-1999; 99US-0134218P.
PR      14-MAY-1999; 99US-0134219P.
PR      14-MAY-1999; 99US-0134221P.
PR      14-MAY-1999; 99US-0134370P.
PR      18-MAY-1999; 99US-0134768P.
PR      19-MAY-1999; 99US-0134941P.
PR      20-MAY-1999; 99US-0135124P.
PR      21-MAY-1999; 99US-0135353P.
PR      24-MAY-1999; 99US-0135629P.
PR      25-MAY-1999; 99US-0136021P.
PR      27-MAY-1999; 99US-0136392P.
PR      28-MAY-1999; 99US-0136782P.
PR      01-JUN-1999; 99US-0137222P.
PR      03-JUN-1999; 99US-0137528P.
PR      04-JUN-1999; 99US-0137502P.
PR      07-JUN-1999; 99US-0137724P.
PR      08-JUN-1999; 99US-0138094P.
PR      10-JUN-1999; 99US-0138540P.
PR      10-JUN-1999; 99US-0138847P.
PR      14-JUN-1999; 99US-0139119P.
PR      16-JUN-1999; 99US-0139452P.
PR      16-JUN-1999; 99US-0139453P.
PR      17-JUN-1999; 99US-0139492P.
PR      18-JUN-1999; 99US-0139454P.
PR      18-JUN-1999; 99US-0139455P.
PR      18-JUN-1999; 99US-0139456P.
PR      18-JUN-1999; 99US-0139457P.
PR      18-JUN-1999; 99US-0139458P.
PR      18-JUN-1999; 99US-0139459P.
PR      18-JUN-1999; 99US-0139460P.
PR      18-JUN-1999; 99US-0139461P.
PR      18-JUN-1999; 99US-0139462P.
PR      18-JUN-1999; 99US-0139463P.
PR      18-JUN-1999; 99US-0139750P.
PR      18-JUN-1999; 99US-0139763P.
PR      21-JUN-1999; 99US-0139817P.
PR      22-JUN-1999; 99US-0139899P.
PR      23-JUN-1999; 99US-0140353P.
PR      23-JUN-1999; 99US-0140354P.
PR      24-JUN-1999; 99US-0140695P.
PR      28-JUN-1999; 99US-0140821P.
PR      29-JUN-1999; 99US-0140991P.
PR      30-JUN-1999; 99US-0141287P.
PR      01-JUL-1999; 99US-0141842P.
PR      01-JUL-1999; 99US-0142154P.
PR      02-JUL-1999; 99US-0142055P.
PR      06-JUL-1999; 99US-0142390P.
PR      08-JUL-1999; 99US-0142803P.
PR      09-JUL-1999; 99US-0142920P.
PR      12-JUL-1999; 99US-0142977P.
PR      13-JUL-1999; 99US-0143542P.
PR      14-JUL-1999; 99US-0143624P.
PR      15-JUL-1999; 99US-0144005P.
PR      16-JUL-1999; 99US-0144085P.
PR      16-JUL-1999; 99US-0144086P.
PR      19-JUL-1999; 99US-0144325P.
PR      19-JUL-1999; 99US-0144331P.
PR      19-JUL-1999; 99US-0144332P.
PR      19-JUL-1999; 99US-0144333P.
PR      19-JUL-1999; 99US-0144334P.
PR      19-JUL-1999; 99US-0144335P.
PR      20-JUL-1999; 99US-0144352P.
PR      20-JUL-1999; 99US-0144632P.
PR      20-JUL-1999; 99US-0144884P.
PR      21-JUL-1999; 99US-0144814P.
PR      21-JUL-1999; 99US-0145086P.
PR      21-JUL-1999; 99US-0145088P.
PR      22-JUL-1999; 99US-0145085P.
PR      22-JUL-1999; 99US-0145087P.
PR      22-JUL-1999; 99US-0145089P.
PR      22-JUL-1999; 99US-0145192P.
PR      23-JUL-1999; 99US-0145145P.
PR      23-JUL-1999; 99US-0145218P.
PR      23-JUL-1999; 99US-0145224P.
PR      26-JUL-1999; 99US-0145276P.
PR      27-JUL-1999; 99US-0145913P.
PR      27-JUL-1999; 99US-0145918P.
PR      27-JUL-1999; 99US-0145919P.
PR      28-JUL-1999; 99US-0145951P.
PR      02-AUG-1999; 99US-0146386P.
PR      02-AUG-1999; 99US-0146388P.
PR      02-AUG-1999; 99US-0146389P.
PR      03-AUG-1999; 99US-0147038P.
PR      04-AUG-1999; 99US-0147204P.
PR      04-AUG-1999; 99US-0147302P.
PR      05-AUG-1999; 99US-0147192P.
PR      05-AUG-1999; 99US-0147260P.
PR      06-AUG-1999; 99US-0147303P.
PR      06-AUG-1999; 99US-0147416P.
PR      09-AUG-1999; 99US-0147493P.
PR      09-AUG-1999; 99US-0147933P.
PR      10-AUG-1999; 99US-0148171P.
PR      11-AUG-1999; 99US-0148319P.
PR      12-AUG-1999; 99US-0148341P.
PR      13-AUG-1999; 99US-0148565P.
```

PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 21-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142290P.  
PR 09-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143342P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.

PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 23-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 25-AUG-1999; 99US-0149930P.  
PR 26-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 73.5%; Score 108; DB 3; Length 400;  
Best Local Similarity 67.9%; Pred. No. 8e-09; 3; Mismatches 6; Indels 0; Gaps 0;  
Matches 19; Conservative

```
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 06-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148117P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0156559P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.

PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 73.5%; Score 108; DB 3; Length 395;
Best Local Similarity 67.9%; Pred. No. 7, 9e-09;
Matches 19; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 SASGLALYPTNRNKNRMSAMTPEID 28
   |||:|||||:|||||:
Db 285 STSGVTLFPTNRNKNRMSMTPTDED 312

RESULT 12
AAG31317
ID AAG31317 standard; protein; 400 AA.
XX
AC AAG31317;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 37590.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
```

CC for herbicidally active compounds, comprising aligning and comparing  
CC nucleic acid or amino acid sequences from plant with nucleic acid or  
CC amino acid sequences from non-plant organisms using suitable search  
CC parameters, where plant sequences having an E-value greater by a factor  
CC of 3 than the E-value of most similar non-plant sequences are selected.  
CC The polypeptides or nucleic acids encoding them are useful for  
CC identifying modulators. The identified modulators are useful as  
CC herbicides

XX Sequence 515 AA;

SQ Query Match 88.4%; Score 130; DB 5; Length 515;

Best Local Similarity 66.7%; Pred. No. 2.6e-12;

Matches 28; Conservative 0; Mismatches 0; Indels 14; Gaps 1;

'QY 1 SASGLALLPTNRN-----KWDNRMSAMPEID 28

Db 391 SASGLALLPTNRNKYLLFIHYLQEPKWDNRMSAMPEID 432

RESULT 11

AAG31318

ID AAG31318 standard; protein; 395 AA.

XX AC AAG31318;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 37591.

XX KW Protein identification; signal transduction pathway; metabolic pathway;

XX KW hybridisation assay; genetic mapping; gene expression control; promoter;

XX KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EF1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-00301439.

XX PR 25-FEB-1999; 99US-0121825P.

XX PR 05-MAR-1999; 99US-0123180P.

XX PR 09-MAR-1999; 99US-0123548P.

XX PR 23-MAR-1999; 99US-0125788P.

XX PR 25-MAR-1999; 99US-0126264P.

XX PR 29-MAR-1999; 99US-0126785P.

XX PR 01-APR-1999; 99US-0127462P.

XX PR 06-APR-1999; 99US-0128234P.

XX PR 08-APR-1999; 99US-0128714P.

XX PR 16-APR-1999; 99US-0129845P.

XX PR 19-APR-1999; 99US-0130077P.

XX PR 21-APR-1999; 99US-0130449P.

XX PR 23-APR-1999; 99US-0130510P.

XX PR 28-APR-1999; 99US-0130891P.

XX PR 30-APR-1999; 99US-0131449P.

XX PR 30-APR-1999; 99US-0132048P.

XX PR 04-MAY-1999; 99US-0132407P.

XX PR 05-MAY-1999; 99US-0132484P.

XX PR 06-MAY-1999; 99US-0132485P.

XX PR 06-MAY-1999; 99US-0132486P.

XX PR 07-MAY-1999; 99US-0132487P.

XX PR 11-MAY-1999; 99US-0132863P.

XX PR 14-MAY-1999; 99US-0134218P.

XX PR 14-MAY-1999; 99US-0134219P.

XX PR 14-MAY-1999; 99US-0134221P.

XX PR 18-MAY-1999; 99US-0134370P.

XX PR 18-MAY-1999; 99US-0134768P.

XX PR 20-MAY-1999; 99US-0134941P.

XX PR 20-MAY-1999; 99US-0135124P.

XX PR 21-MAY-1999; 99US-0135353P.

PR 24-MAY-1999; 99US-0135629P.

PR 25-MAY-1999; 99US-0136021P.

PR 27-MAY-1999; 99US-0136392P.

PR 28-MAY-1999; 99US-0136782P.

PR 01-JUN-1999; 99US-0137222P.

PR 03-JUN-1999; 99US-0137528P.

PR 04-JUN-1999; 99US-0137502P.

PR 07-JUN-1999; 99US-0137724P.

PR 08-JUN-1999; 99US-0138094P.

PR 10-JUN-1999; 99US-0138540P.

PR 10-JUN-1999; 99US-0138847P.

PR 14-JUN-1999; 99US-0139119P.

PR 16-JUN-1999; 99US-0139452P.

PR 16-JUN-1999; 99US-0139453P.

PR 17-JUN-1999; 99US-0139492P.

PR 18-JUN-1999; 99US-0139454P.

PR 18-JUN-1999; 99US-0139455P.

PR 18-JUN-1999; 99US-0139456P.

PR 18-JUN-1999; 99US-0139457P.

PR 18-JUN-1999; 99US-0139458P.

PR 18-JUN-1999; 99US-0139459P.

PR 18-JUN-1999; 99US-0139460P.

PR 18-JUN-1999; 99US-0139461P.

PR 18-JUN-1999; 99US-0139462P.

PR 18-JUN-1999; 99US-0139463P.

PR 18-JUN-1999; 99US-0139750P.

PR 18-JUN-1999; 99US-0139763P.

PR 21-JUN-1999; 99US-0139817P.

PR 22-JUN-1999; 99US-0139899P.

PR 23-JUN-1999; 99US-0140153P.

PR 23-JUN-1999; 99US-0140354P.

PR 24-JUN-1999; 99US-0140695P.

PR 28-JUN-1999; 99US-0140823P.

PR 29-JUN-1999; 99US-0140991P.

PR 30-JUN-1999; 99US-0141287P.

PR 01-JUL-1999; 99US-0141842P.

PR 01-JUL-1999; 99US-0142154P.

PR 02-JUL-1999; 99US-0142055P.

PR 06-JUL-1999; 99US-0142390P.

PR 08-JUL-1999; 99US-0142803P.

PR 09-JUL-1999; 99US-0142920P.

PR 12-JUL-1999; 99US-0142977P.

PR 13-JUL-1999; 99US-0143542P.

PR 14-JUL-1999; 99US-0143624P.

PR 15-JUL-1999; 99US-0144005P.

PR 16-JUL-1999; 99US-0144085P.

PR 16-JUL-1999; 99US-0144086P.

PR 19-JUL-1999; 99US-0144325P.

PR 19-JUL-1999; 99US-0144331P.

PR 19-JUL-1999; 99US-0144332P.

PR 19-JUL-1999; 99US-0144333P.

PR 19-JUL-1999; 99US-0144334P.

PR 19-JUL-1999; 99US-0144335P.

PR 20-JUL-1999; 99US-0144352P.

PR 20-JUL-1999; 99US-0144632P.

PR 20-JUL-1999; 99US-0144884P.

PR 21-JUL-1999; 99US-0144814P.

PR 21-JUL-1999; 99US-0145086P.

PR 21-JUL-1999; 99US-0145088P.

PR 22-JUL-1999; 99US-0145085P.

PR 22-JUL-1999; 99US-0145087P.

PR 22-JUL-1999; 99US-0145089P.

PR 22-JUL-1999; 99US-0145192P.

PR 23-JUL-1999; 99US-0145145P.

PR 23-JUL-1999; 99US-0145218P.

PR 23-JUL-1999; 99US-0145224P.

PR 26-JUL-1999; 99US-0145276P.

PR 27-JUL-1999; 99US-0145913P.

PR 27-JUL-1999; 99US-0145918P.

PR 27-JUL-1999; 99US-0145919P.

PR 28-JUL-1999; 99US-0145951P.

PR 02-AUG-1999; 99US-0146386P.

PR 02-AUG-1999; 99US-0146388P.





PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 10-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 26-JUL-1999; 99US-0145224P.  
PR 27-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 28-JUL-1999; 99US-0145919P.  
PR 02-AUG-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 25-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.

PR 01-JUL-1999; 99US-01411842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142820P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143342P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 23-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0147260P.  
PR 02-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.

PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 18-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 88.4%; Score 130; DB 3; Length 501;  
Best Local Similarity 66.7%; Pred. No. 2.5e-12;  
Matches 28; Conservative 0; Mismatches 0; Indels 14; Gaps 1;

QY 1 SASGLALYPTNRN-----KWDNRMSAMTPEID 28  
DB 377 SASGLALYPTNRNKYILLFIHYLQEPKWDNRMSAMTPEID 418

RESULT 8  
AAG31941  
ID AAG31941 standard; protein; 509 AA.

XX  
AC AAG31941;  
XX  
DT 17-OCT-2000 (first entry)  
XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 38445.

KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX

KW metal deficiency; nutrition; cytokinin oxidase; enzyme.  
 XX Arabidopsis thaliana.  
 XX WO2004038027-A1.  
 PN 06-MAY-2004.  
 PD 24-OCT-2003; 2003WO-BP012051.  
 XX 24-OCT-2002; 2002EP-00079481.  
 XX (CROP-) CROPDESIGN NV.  
 PA Van Camp W;  
 PI WPI; 2004-375913/35.  
 XX N-PSDB; ADO06498.  
 DR Method for bioremediation, useful for removing contaminants or metals, by  
 PT decreasing cytokinin availability in plants, and cultivating plant on  
 PT substrate comprising one or more contaminants.  
 XX Claim 8; Page 55-56; 61pp; English.  
 XX The present invention relates to a method for bioremediation, which  
 CC involves decreasing cytokinin availability in a plant relative to a  
 CC corresponding wild type plant, and cultivating the plant on a substrate  
 CC comprising one or more contaminants, or cultivating a plant having  
 CC lowered availability of cytokinin relative to corresponding wild type  
 CC plants, on a substrate, which is to be treated. The method is useful for  
 CC bioremediation, for concentrating contaminants in a plant, where the  
 CC plant has a higher concentration of contaminants compared to a  
 CC corresponding wild type plant. The metal contaminants include aluminum,  
 CC americium, antimony, arsenic, barium, beryllium, bismuth, cadmium,  
 CC caesium, cerium, chromium, copper, gallium, germanium, gold, indium,  
 CC iridium, iron, lead, manganese, mercury, molybdenum, neptunium, osmium,  
 CC palladium, platinum, plutonium, radium, rhodium, rubidium,  
 CC ruthenium, scandium, selenium, silver, strontium, technetium, tellurium,  
 CC thallium, tin, tungsten, uranium, vanadium or yttrium, preferably  
 CC cadmium. The plant obtained by the method is useful in bioremediation.  
 CC Transgenic plants with a lower availability of cytokinin are useful in  
 CC the manufacture of a medicament for treatment of disorders arising from  
 CC metal deficiencies, and as a medicament for improving animal or human  
 CC nutrition. The present sequence is a protein shown in the exemplification  
 CC of the invention.  
 XX Sequence 501 AA;  
 SQ  
 Query Match 100.0%; Score 147; DB 8; Length 501;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-15;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SASGLALLPTNRNKNWDRMSAMIPED 28  
 DB 391 SASGLALLPTNRNKNWDRMSAMIPED 418  
 RESULT 7  
 AAG31942  
 ID AAG31942 standard; protein; 501 AA.  
 XX AAG31942;  
 AC  
 XX 17-OCT-2000 (first entry)  
 DT Arabidopsis thaliana protein fragment SEQ ID NO: 38446.  
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 38446.  
 XX Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX Arabidopsis thaliana.  
 OS

XX EP1033405-A2.  
 PN 06-SEP-2000.  
 PD 25-FEB-2000; 2000EP-00301439.  
 PF 25-FEB-1999; 99US-0121825P.  
 XX 05-MAR-1999; 99US-0123180P.  
 PR 09-MAR-1999; 99US-0123548P.  
 PR 23-MAR-1999; 99US-0125788P.  
 PR 25-MAR-1999; 99US-0126264P.  
 PR 29-MAR-1999; 99US-0126785P.  
 PR 01-APR-1999; 99US-0127462P.  
 PR 06-APR-1999; 99US-0128234P.  
 PR 08-APR-1999; 99US-0128714P.  
 PR 16-APR-1999; 99US-0129845P.  
 PR 19-APR-1999; 99US-0130077P.  
 PR 21-APR-1999; 99US-0130498P.  
 PR 23-APR-1999; 99US-0130510P.  
 PR 23-APR-1999; 99US-0130891P.  
 PR 28-APR-1999; 99US-0131449P.  
 PR 30-APR-1999; 99US-0132048P.  
 PR 30-APR-1999; 99US-0132407P.  
 PR 04-MAY-1999; 99US-0132484P.  
 PR 05-MAY-1999; 99US-0132485P.  
 PR 06-MAY-1999; 99US-0132486P.  
 PR 07-MAY-1999; 99US-0132487P.  
 PR 11-MAY-1999; 99US-0132863P.  
 PR 11-MAY-1999; 99US-0134256P.  
 PR 14-MAY-1999; 99US-0134218P.  
 PR 14-MAY-1999; 99US-0134219P.  
 PR 14-MAY-1999; 99US-0134221P.  
 PR 14-MAY-1999; 99US-0134370P.  
 PR 18-MAY-1999; 99US-0134768P.  
 PR 19-MAY-1999; 99US-0134941P.  
 PR 20-MAY-1999; 99US-0135124P.  
 PR 21-MAY-1999; 99US-0135353P.  
 PR 24-MAY-1999; 99US-0135629P.  
 PR 25-MAY-1999; 99US-0136021P.  
 PR 27-MAY-1999; 99US-0136392P.  
 PR 28-MAY-1999; 99US-0136782P.  
 PR 01-JUN-1999; 99US-0137222P.  
 PR 03-JUN-1999; 99US-0137528P.  
 PR 04-JUN-1999; 99US-0137502P.  
 PR 07-JUN-1999; 99US-0137724P.  
 PR 08-JUN-1999; 99US-0138094P.  
 PR 10-JUN-1999; 99US-0138540P.  
 PR 10-JUN-1999; 99US-0138847P.  
 PR 14-JUN-1999; 99US-0139119P.  
 PR 16-JUN-1999; 99US-0139452P.  
 PR 16-JUN-1999; 99US-0139453P.  
 PR 17-JUN-1999; 99US-0139492P.  
 PR 18-JUN-1999; 99US-0139454P.  
 PR 18-JUN-1999; 99US-0139455P.  
 PR 18-JUN-1999; 99US-0139456P.  
 PR 18-JUN-1999; 99US-0139457P.  
 PR 18-JUN-1999; 99US-0139458P.  
 PR 18-JUN-1999; 99US-0139459P.  
 PR 18-JUN-1999; 99US-0139460P.  
 PR 18-JUN-1999; 99US-0139461P.  
 PR 18-JUN-1999; 99US-0139462P.  
 PR 18-JUN-1999; 99US-0139463P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 18-JUN-1999; 99US-0139763P.  
 PR 21-JUN-1999; 99US-0139817P.  
 PR 22-JUN-1999; 99US-0139899P.  
 PR 23-JUN-1999; 99US-0140353P.  
 PR 23-JUN-1999; 99US-0140354P.  
 PR 24-JUN-1999; 99US-0140695P.  
 PR 28-JUN-1999; 99US-0140823P.  
 PR 29-JUN-1999; 99US-0140991P.  
 PR 30-JUN-1999; 99US-0141287P.

Query Match 100.0%; Score 147; DB 5; Length 501;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-15;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SASGLALLYPTRNKNKWNMSAMIPED 28  
 |||||  
 DB 391 SASGLALLYPTRNKNKWNMSAMIPED 418

## RESULT 4

ID ABR63565  
 AC ABR63565; standard; protein; 501 AA.

XX ABR63565;

DT 18-SEP-2003 (first entry)

DE Arabidopsis cytokinin oxidase AtCKX2.

KW Maize; root growth; root geotropism; cytokinin oxidase; seed size;  
 KW embryo size; cotyledon size; transgenic plant; herbicide.

OS Arabidopsis thaliana.

XX WO2003050287-A2.

PN 19-JUN-2003.

PF 10-DEC-2002; 2002WO-EP013990.

PR 10-DEC-2001; 2001US-00014101.

XX (SCHM/) SCHMULLING T.

PA (WERN/) WERNER T.

XX Schmullling T, Werner T;

XX WPI; 2003-541577/51.

XX Stimulating root growth, enhancing lateral or adventitious root formation  
 or altering root geotropism comprises increasing plant cytokinin oxidase  
 levels or other protein or nucleic acid that reduces active cytokinins in  
 a plant.

PS Example 2; Fig 2; 177pp; English.

XX The present invention relates to a method for stimulating root growth or  
 enhancing the formation of lateral or adventitious roots or altering root  
 geotropism, which comprises increasing in a plant or plant part the level  
 of a plant cytokinin oxidase or other protein that reduces the level of  
 active cytokinins in a plant or plant part. Cytokinin oxidase protein and  
 coding sequences from Arabidopsis thaliana are also provided. The method  
 is useful in modifying plant morphological, biochemical and physiological  
 properties, such as in modifying the initiation, stimulation or  
 enhancement of root growth, adventitious root formation, lateral root  
 formation, root geotropism, shoot growth, apical dominance, branching,  
 timing of senescence, timing of flowering, flower formation, seed  
 development and/or seed yield. The present sequence is a protein shown in  
 the invention

XX Sequence 501 AA;

Query Match 100.0%; Score 147; DB 6; Length 501;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-15;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SASGLALLYPTRNKNKWNMSAMIPED 28  
 |||||  
 DB 391 SASGLALLYPTRNKNKWNMSAMIPED 418

## RESULT 5

## ABR63569

ID ABR63569 standard; protein; 501 AA.

XX ABR63569;

XX 18-SEP-2003 (first entry)

DE Arabidopsis cytokinin oxidase-like protein 2.

KW Maize; root growth; root geotropism; cytokinin oxidase; seed size;  
 KW embryo size; cotyledon size; transgenic plant; herbicide.

XX Arabidopsis thaliana.

XX WO2003050287-A2.

XX 19-JUN-2003.

XX 10-DEC-2002; 2002WO-EP013990.

PR 10-DEC-2001; 2001US-00014101.

PA (SCHM/) SCHMULLING T.

PA (WERN/) WERNER T.

XX Schmullling T, Werner T;

XX WPI; 2003-541577/51.

DR N-PSDB; ACC65276.

XX Stimulating root growth, enhancing lateral or adventitious root formation  
 or altering root geotropism comprises increasing plant cytokinin oxidase  
 levels or other protein or nucleic acid that reduces active cytokinins in  
 a plant.

PS Claim 19; Page 146-148; 177pp; English.

XX The present invention relates to a method for stimulating root growth or  
 enhancing the formation of lateral or adventitious roots or altering root  
 geotropism, which comprises increasing in a plant or plant part the level  
 of a plant cytokinin oxidase or other protein that reduces the level of  
 active cytokinins in a plant or plant part. Cytokinin oxidase protein and  
 coding sequences from Arabidopsis thaliana are also provided. The method  
 is useful in modifying plant morphological, biochemical and physiological  
 properties, such as in modifying the initiation, stimulation or  
 enhancement of root growth, adventitious root formation, lateral root  
 formation, root geotropism, shoot growth, apical dominance, branching,  
 timing of senescence, timing of flowering, flower formation, seed  
 development and/or seed yield. The present sequence is a protein shown in  
 the invention

XX Sequence 501 AA;

Query Match 100.0%; Score 147; DB 6; Length 501;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-15;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SASGLALLYPTRNKNKWNMSAMIPED 28  
 |||||  
 DB 391 SASGLALLYPTRNKNKWNMSAMIPED 418

## RESULT 6

ADO06499

ID ADO06499 standard; protein; 501 AA.

XX ADO06499;

XX 29-JUL-2004 (first entry)

DE A thaliana cytokinin oxidase AtCKX2 protein.

XX bioremediation; AtCKX2; cytokinin availability; contaminant;

CC geotropism, leading to an increase in yield; and for screening growth  
 CC promoting chemical of herbicides. (I) is useful for increasing the size  
 CC of the root meristem; increasing root size; increasing the size of the  
 CC shoot meristem; delaying leaf senescence and altering leaf senescence;  
 CC increasing leaf thickness; reducing or increasing the vessel size;  
 CC inducing parthenocarp; improving standability of the seedlings;  
 CC increasing parthenocarp; and for improving lodging resistance. Antibody (III)  
 CC to (II) is useful for identifying and obtaining proteins interacting with  
 CC (II) comprising a screening assay, preferably a two-hybrid screening  
 CC assay. AAU81967-AAU81974 represent Arabidopsis thaliana cytokinin oxidase  
 CC amino acid sequences  
 XX  
 SQ Sequence 28 AA;

Query Match 100.0%; Score 147; DB 5; Length 28;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-16;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SASGLALYPTNRNKNWDRMSAMIPED 28  
 |||||  
 Db 1 SASGLALYPTNRNKNWDRMSAMIPED 28

RESULT 2  
 ABR63574  
 ID ABR63574 standard; peptide; 28 AA.  
 XX  
 AC ABR63574;  
 XX  
 DT 18-SEP-2003 (first entry)  
 XX  
 DE Arabidopsis cytokinin oxidase-like protein 2 fragment.  
 XX  
 KW Maize; root growth; root geotropism; cytokinin oxidase; seed size;  
 KW embryo size; cotyledon size; transgenic plant; herbicide.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN WO2003050287-A2.  
 XX  
 PD 19-JUN-2003.  
 XX  
 PF 10-DEC-2002; 2002WO-EP013990.  
 XX  
 PR 10-DEC-2001; 2001US-00014101.  
 XX

(SCHM/) SCHMULLING T.  
 PA (WERN/) WERNER T.  
 XX  
 PI Schullling T, Werner T;  
 XX  
 DR WPI; 2003-541577/51.  
 XX  
 ST Stimulating root growth, enhancing lateral or adventitious root formation  
 PT or altering root geotropism comprises increasing plant cytokinin oxidase  
 PT levels or other protein or nucleic acid that reduces active cytokinins in  
 PT a plant.  
 XX  
 PS Claim 3; Page 172; 177pp; English.  
 XX

CC The present invention relates to a method for stimulating root growth or  
 CC enhancing the formation of lateral or adventitious roots or altering root  
 CC geotropism, which comprises increasing in a plant or plant part the level  
 CC of a plant cytokinin oxidase or other protein that reduces the level of  
 CC active cytokinins in a plant or plant part. Cytokinin oxidase protein and  
 CC coding sequences from Arabidopsis thaliana are also provided. The method  
 CC is useful in modifying plant morphological, biochemical and physiological  
 CC properties, such as in modifying the initiation, stimulation or  
 CC enhancement of root growth, adventitious root formation, lateral root  
 CC formation, root geotropism, shoot growth, apical dominance, branching,  
 CC timing of senescence, timing of flowering, flower formation, seed  
 CC development and/or seed yield. The present sequence is a protein shown in  
 CC the invention

XX SQ Sequence 28 AA;  
 Query Match 100.0%; Score 147; DB 6; Length 28;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-16;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SASGLALYPTNRNKNWDRMSAMIPED 28  
 |||||  
 Db 1 SASGLALYPTNRNKNWDRMSAMIPED 28

RESULT 3  
 AAU81968  
 ID AAU81968 standard; protein; 501 AA.  
 XX  
 AC AAU81968;  
 XX  
 DT 09-APR-2002 (first entry)  
 XX  
 DE A. thaliana cytokinin oxidase AtCKX2.  
 XX  
 KW Arabidopsis thaliana; plant; cytokinin oxidase; transgenic plant;  
 KW root growth; lateral root; adventitious root; root geotropism; herbicide;  
 KW root meristem; shoot meristem; leaf senescence; parthenocarp.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN WO200196580-A2.  
 XX  
 PD 20-DEC-2001.  
 XX  
 PF 18-JUN-2001; 2001WO-EP006833.  
 XX  
 PR 16-JUN-2000; 2000EP-00870132.  
 PR 27-DEC-2000; 2000US-0258415P.  
 PR 16-MAR-2001; 2001EP-00870053.  
 XX  
 PA (SCHM/) SCHMULLING T.  
 PA (WERN/) WERNER T.  
 XX  
 PI Schullling T, Werner T;  
 XX  
 DR WPI; 2002-130736/17.  
 DR N-PSDB; ABK28607.  
 XX  
 PT Polynucleotide encoding novel plant protein having cytokinin oxidase  
 PT activity and the protein useful for stimulating root growth, enhancing  
 PT the formation of lateral or adventitious roots, altering root geotropism.  
 XX  
 PS Claim 12; Page 129-131; 154pp; English.  
 XX

CC The invention relates to an isolated polynucleotide (I) encoding a novel  
 CC plant protein (II) having cytokinin oxidase activity. (I) is useful for  
 CC production of transgenic plants, plant cells or tissues; for production  
 CC of altered plants, plant cell or tissues; and for effecting the  
 CC expression of (II) where (I) is operably linked to one or more control  
 CC sequences. The methods further comprises regenerating a plant from the  
 CC plant cell. (I) and (II) are useful for stimulating root growth;  
 CC enhancing the formation of lateral or adventitious roots; altering root  
 CC geotropism, leading to an increase in yield; and for screening growth  
 CC promoting chemical of herbicides. (I) is useful for increasing the size  
 CC of the root meristem; increasing root size; increasing the size of the  
 CC shoot meristem; delaying leaf senescence and altering leaf senescence;  
 CC increasing leaf thickness; reducing or increasing the vessel size;  
 CC inducing parthenocarp; improving standability of the seedlings;  
 CC increasing branching and for improving lodging resistance. Antibody (III)  
 CC to (II) is useful for identifying and obtaining proteins interacting with  
 CC (II) comprising a screening assay, preferably a two-hybrid screening  
 CC assay. AAU81967-AAU81974 represent Arabidopsis thaliana cytokinin oxidase  
 CC amino acid sequences  
 XX  
 SQ Sequence 501 AA;

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 22, 2005, 22:18:46 ; Search time 166 Seconds  
(without alignments)  
65.237 Million cell updates/sec

Title: US-10-014-101B-32

Perfect score: 147  
Sequence: 1 SASGLALLYPTNRKWNRSAMIPSID 28

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	147	100.0	28	5 AAU81973	Aau81973 A. thalia
2	147	100.0	28	6 ABR63574	Abr63574 Arabidops
3	147	100.0	501	5 AAU81968	Aau81968 A. thalia
4	147	100.0	501	6 ABR63565	Abr63565 Arabidops
5	147	100.0	501	6 ABR63569	Abr63569 Arabidops
6	147	100.0	501	8 ADO06499	Ado06499 A thalian
7	130	88.4	501	3 AAG31942	Aag31942 Arabidops
8	130	88.4	509	3 AAG31941	Aag31941 Arabidops
9	130	88.4	515	3 AAG31940	Aag31940 Arabidops
10	130	88.4	515	5 ABB91811	Abb91811 Arabidops
11	108	73.5	395	3 AAG31318	Aag31318 Arabidops
12	108	73.5	400	3 AAG31317	Aag31317 Arabidops
13	108	73.5	524	3 AAG31316	Aag31316 Arabidops
14	108	73.5	524	5 ABB93161	Abb93161 Arabidops
15	108	73.5	524	5 AAU81970	Aau81970 A. thalia
16	108	73.5	524	6 ABR63567	Abr63567 Arabidops
17	108	73.5	524	6 ABR63571	Abr63571 Arabidops
18	96	65.3	523	5 ABB93847	Abb93847 Arabidops
19	96	65.3	523	5 AAU81969	Aau81969 A. thalia
20	96	65.3	523	6 ABR63566	Abr63566 Arabidops
21	96	65.3	523	6 ABR63570	Abr63570 Arabidops
22	87	59.2	563	8 ADP03326	Adp03326 Rice grai
23	87	59.2	565	8 ADP03323	Adp03323 Rice grai
24	80	54.4	504	5 ABB92780	Abb92780 Arabidops
25	80	54.4	504	5 AAU81972	Aau81972 A. thalia

26	80	54.4	504	6 ABR63573	Abr63573 Arabidops
27	77	52.4	536	5 ABB91608	Abb91608 Arabidops
28	77	52.4	536	5 AAU81971	Aau81971 A. thalia
29	77	52.4	536	6 ABR63572	Abr63572 Arabidops
30	77	52.4	539	5 AAU81974	Aau81974 A. thalia
31	77	52.4	539	6 ABR63575	Abr63575 Arabidops
32	63	42.9	534	2 AAU81970	Aau81970 A. thalia
33	63	42.9	534	6 ABR63563	Abr63563 Maize cyt
34	63	42.9	535	3 AAB30691	Aab30691 Amino aci
35	63	42.9	611	5 AAE13725	Aae13725 pUE-7 pla
36	58	39.5	575	5 ABB92071	Abb92071 Arabidops
37	58	39.5	575	5 AAU81967	Aau81967 A. thalia
38	58	39.5	575	6 ABR63568	Abr63568 Arabidops
39	58	39.5	575	6 ABR63564	Abr63564 Arabidops
40	58	39.5	575	8 ADO06497	Ado06497 A thalian
41	52	35.4	269	3 AAB53146	Aab53146 Macaca mu
42	50	34.0	1322	8 ADI42574	Adi42574 Plant tra
43	50	34.0	1322	8 ADO62380	Ado62380 Transcrip
44	49	33.3	257	2 AAU81977	Aau81977 Kaposi's
45	49	33.3	280	6 ABU49494	Abu49494 Protein e

## ALIGNMENTS

## RESULT 1

AAU81973  
ID AAU81973 standard; peptide; 28 AA.  
XX  
AC AAU81973;  
XX  
DT 09-APR-2002 (first entry)  
XX  
DE A. thaliana cytokinin oxidase AtCKX2, peptide.  
XX  
KW Arabidopsis thaliana; plant; cytokinin oxidase; transgenic plant;  
KW root growth; lateral root; adventitious root; root geotropism; herbicide;  
KW root meristem; shoot meristem; leaf senescence; parthenocary.  
XX  
OS Arabidopsis thaliana.  
XX  
PN WO200196580-A2.  
XX  
PD 20-DEC-2001.  
XX  
PF 18-JUN-2001; 2001WO-EP006833.  
XX  
PR 16-JUN-2000; 2000EP-00870132.  
PR 27-DEC-2000; 2000US-0258415P.  
PR 16-MAR-2001; 2001EP-00870053.  
(SCHM/) SCHMULLING T.  
PA (WERN/) WERNER T.  
XX  
PI Schmulling T, Werner T;  
XX WPI; 2002-130736/17.  
DR N-PSDB; ABK28630.  
DR  
PT Polynucleotide encoding novel plant protein having cytokinin oxidase  
PT activity and the protein useful for stimulating root growth, enhancing  
PT the formation of lateral or adventitious roots, altering root geotropism.  
PS Claim 3; Page 150; 154pp; English.  
XX  
CC The invention relates to an isolated polynucleotide (I) encoding a novel  
CC plant protein (II) having cytokinin oxidase activity. (I) is useful for  
CC production of transgenic plants, plant cells or tissues; for production  
CC of altered plants, plant cell or tissues; and for effecting the  
CC expression of (II), where (I) is operably linked to one or more control  
CC sequences. The methods further comprises regenerating a plant from the  
CC plant cell. (I) and (II) are useful for stimulating root growth; altering  
CC enhancing the formation of lateral or adventitious roots; altering root

```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221,017B
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P1182
; FILING DATE: 31-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P1546
; FILING DATE: 30-JAN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P2911
; FILING DATE: 09-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU98/01023
; FILING DATE: 10-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Monroy, Gladys H
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 27340-20021.00
; TELEPHONE: 650-813-5600
; TELEFAX: 650-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 1014:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2601 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: UNKNOWN
; ORIGINAL SOURCE:
; ORGANISM: PORPHYROMONAS GINGIVALIS
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1...2601
;
; US-09-221-017B-1014
;
; Alignment Scores:
; Pred. No.: 272 Length: 2601
; Score: 49.00 Matches: 8
; Percent Similarity: 64.71% Conservative: 3
; Best Local Similarity: 47.06% Mismatches: 6
; Query Match: 33.33% Indels: 0
; DB: 3 Gaps: 0
;
; US-10-014-101B-32 (1-28) x US-09-221-017B-1014 (1-2601)
;
; Qy 1 SerAlaSerGlyLeuAlaLeuTyrProThrAsnArgAsnLysTrpAsp 17
; Db 2230 TCCACATCTATCGTATCTTATTCTCTGCAACACAGAGAGAGTGGGAT 2180
;
; RESULT 40
; US-08-770-379-20/c
; Sequence 20, Application US/08770379
; Patent No. 5849564
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Edelman, Isidore S.
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED
; HERPESVIRUS, DNA ENCODING SAME AND USES THEREOF
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
;
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/770,379
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 52342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32207 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
;
; US-08-770-379-20
;
; Alignment Scores:
; Pred. No.: 7,29e+03 Length: 32207
; Score: 49.00 Matches: 10
; Percent Similarity: 55.56% Conservative: 5
; Best Local Similarity: 37.04% Mismatches: 12
; Query Match: 33.33% Indels: 0
; DB: 2 Gaps: 0
;
; US-10-014-101B-32 (1-28) x US-08-770-379-20 (1-32207)
;
; Qy 1 SerAlaSerGlyLeuAlaLeuTyrProThrAsnArgAsnLysTrpAspArgMet 20
; Db 17675 AGCGCTGCAGGCTGCGCTGTGGTTCTCTGCCAACGTCATTCCGCGAGGATACCCACTCG 17616
;
; Qy 21 SerAlaMetIleProGluIle 27
; Db 17615 GGTGGGGTAGTCTCTCAGCTG 17595
;
; Search completed: February 18, 2005, 14:16:23
; Job time : 557 secs
```

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 18, 2005, 12:07:27 ; Search time 289 Seconds  
(without alignments)  
572.643 Million cell updates/sec

Title: US-10-014-101b-32

Perfect score: 147

Sequence: 1 SASGLALLPTNRKWDNRMSAMIPRID 28

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 5384158 seqs, 2955248155 residues

Total number of hits satisfying chosen parameters: 10768316

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10014101/runat\_16022005\_075936\_8410/app\_query.fasta\_1.199  
-DB=Published Applications NA -QFMT=fastp -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=40 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10014101\_@CGN 1 1 354 @runat\_16022005\_075936\_8410  
-NCPU=6 -ICPU=3 -NO MMAP -LARGQUEURY -NEG SCORES=0 -WAIT -DSBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq:  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:  
18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq:  
19: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:  
20: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:  
21: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:  
22: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	147	100.0	84	14	US-10-014-101-31	Sequence 31, Appl
2	147	100.0	84	17	US-10-311-453-31	Sequence 31, Appl
3	147	100.0	1506	14	US-10-014-101-26	Sequence 26, Appl
4	147	100.0	1506	16	US-10-326-184-8	Sequence 8, Appl
5	147	100.0	1506	17	US-10-311-453-26	Sequence 26, Appl
6	108	73.5	1575	9	US-09-938-842A-2089	Sequence 2089, Ap
7	108	73.5	1575	11	US-09-938-842A-2089	Sequence 2089, Ap
8	108	73.5	1575	14	US-10-014-101-28	Sequence 28, Appl
9	108	73.5	1575	16	US-10-326-184-10	Sequence 10, Appl
10	108	73.5	1575	17	US-10-311-453-28	Sequence 28, Appl
11	108	73.5	1655	16	US-10-326-184-20	Sequence 20, Appl
12	96.5	65.6	2991	14	US-10-014-101-3	Sequence 3, Appl
13	96.5	65.6	2991	17	US-10-311-453-3	Sequence 3, Appl
14	96	65.3	1572	14	US-10-014-101-27	Sequence 27, Appl
15	96	65.3	1572	16	US-10-326-184-9	Sequence 9, Appl
16	96	65.3	1572	17	US-10-311-453-27	Sequence 27, Appl
17	91	61.9	1141	17	US-10-425-114-10886	Sequence 10886, A
18	91	61.9	1142	17	US-10-424-599-1725	Sequence 1725, Ap
19	91	61.9	1279	17	US-10-425-114-8021	Sequence 8021, Ap
20	91	61.9	1587	17	US-10-424-599-1724	Sequence 1724, Ap
21	88	59.9	2054	18	US-10-437-963-46009	Sequence 46009, A
22	87	59.2	1677	18	US-10-437-963-43981	Sequence 43981, A
23	86	58.5	1617	18	US-10-817-483-28	Sequence 28, Appl
24	85	57.8	1023	17	US-10-425-114-10804	Sequence 10804, A
25	85	57.8	1045	17	US-10-424-599-19583	Sequence 19583, A
26	85	57.8	1673	18	US-10-437-963-83827	Sequence 83827, A
27	83	56.5	1042	17	US-10-425-114-15562	Sequence 15562, A
28	83	56.5	1560	18	US-10-817-483-26	Sequence 26, Appl
29	83	56.5	1670	18	US-10-437-963-70571	Sequence 70571, A
30	83	56.5	1887	18	US-10-425-115-35578	Sequence 35578, A
31	83	56.5	2782	14	US-10-014-101-7	Sequence 7, Appl
32	83	56.5	2782	17	US-10-311-453-7	Sequence 7, Appl
33	82	55.8	816	17	US-10-424-599-62631	Sequence 62631, A
34	81	55.1	1605	18	US-10-437-963-38170	Sequence 38170, A
35	80	54.4	1515	14	US-10-014-101-30	Sequence 30, Appl
36	80	54.4	1515	17	US-10-311-453-30	Sequence 30, Appl
37	79	53.7	482	18	US-10-021-323-9662	Sequence 9662, Ap
38	79	53.7	484	18	US-10-021-323-10415	Sequence 10415, A
39	79	53.7	788	18	US-10-767-795-2077	Sequence 2077, Ap
40	77	52.4	1611	14	US-10-014-101-29	Sequence 29, Appl
41	77	52.4	1611	17	US-10-311-453-29	Sequence 29, Appl
42	77	52.4	1620	14	US-10-014-101-34	Sequence 34, Appl
43	77	52.4	1620	17	US-10-311-453-34	Sequence 34, Appl
44	77	52.4	1623	16	US-10-326-184-12	Sequence 12, Appl
45	76	51.7	675	17	US-10-425-114-32791	Sequence 32791, A

ALIGNMENTS

RESULT 1

US-10-014-101-31  
; Sequence 31, Application US/10014101  
; Publication No. US20030074698A1  
; GENERAL INFORMATION:  
; APPLICANT: Schmulling, Thomas  
; APPLICANT: Werner, Tom 8  
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and  
; TITLE OF INVENTION: physiology  
; FILE REFERENCE: 1195-2  
; CURRENT FILING DATE: 2001-12-10  
; PRIOR FILING DATE: 2001-06-16  
; PRIOR APPLICATION NUMBER: PCT/EP01/06833  
; PRIOR FILING DATE: 2000-06-16  
; PRIOR APPLICATION NUMBER: EP 00870132.8  
; PRIOR FILING DATE: 2000-06-16  
; PRIOR APPLICATION NUMBER: US 60/258,415  
; PRIOR FILING DATE: 2000-12-27  
; PRIOR APPLICATION NUMBER: EP 01870053.4  
; PRIOR FILING DATE: 2001-03-16

```
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 84
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-014-101-31

Alignment Scores:
Pred. No.:      3,09e-17      Length:      84
Score:          147.00      Matches:      28
Percent Similarity: 100.00%      Conservatives: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              14      Gaps:      0

US-10-014-101B-32 (1-28) x US-10-014-101-31 (1-84)

QY      1 SerAlaSerGlyLeuAlaLeuLeuTyrProThrAsnArgAsnLysTrpAspAsnArgMet 20
Db      1 TCAGCTTCGGGACTCGCTCTCTCTATCCAAACCGGAATAAATGGGACATCGTATG 60

QY      21 SerAlaMetIleProGluIleAsp 28
Db      61 TCGCGCATGATACCAAGATCGAT 84

RESULT 2
US-10-311-453-31
; Sequence 31, Application US/10311453
; Publication No. US20040031073A1
; GENERAL INFORMATION:
; APPLICANT: Schmulling, Thomas
; APPLICANT: Werner, Tom s
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and
; FILE REFERENCE: 1226-4
; CURRENT APPLICATION NUMBER: US/10/311,453
; PRIOR FILING DATE: 2002-12-16
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: EP 00870132.8
; PRIOR FILING DATE: 2000-06-16
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: EP 01870053.4
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 84
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-311-453-31

Alignment Scores:
Pred. No.:      3,09e-17      Length:      84
Score:          147.00      Matches:      28
Percent Similarity: 100.00%      Conservatives: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              17      Gaps:      0

US-10-014-101B-32 (1-28) x US-10-311-453-31 (1-84)

QY      1 SerAlaSerGlyLeuAlaLeuLeuTyrProThrAsnArgAsnLysTrpAspAsnArgMet 20
Db      1 TCAGCTTCGGGACTCGCTCTCTCTATCCAAACCGGAATAAATGGGACATCGTATG 60

QY      21 SerAlaMetIleProGluIleAsp 28
Db      61 TCGCGCATGATACCAAGATCGAT 84

RESULT 3
US-10-014-101-26
; Sequence, 26, Application US/10014101
; Publication No. US20030074698A1
; GENERAL INFORMATION:
; APPLICANT: Schmulling, Thomas
; APPLICANT: Werner, Tom s
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and
; FILE REFERENCE: 1195-2
; CURRENT APPLICATION NUMBER: US/10/014,101
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: PCT/EP01/06833
; PRIOR FILING DATE: 2001-06-16
; PRIOR APPLICATION NUMBER: EP 00870132.8
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/258,415
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: EP 01870053.4
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 1506
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-014-101-26

Alignment Scores:
Pred. No.:      1,2e-15      Length:      1506
Score:          147.00      Matches:      28
Percent Similarity: 100.00%      Conservatives: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              14      Gaps:      0

US-10-014-101B-32 (1-28) x US-10-014-101-26 (1-1506)

QY      1 SerAlaSerGlyLeuAlaLeuLeuTyrProThrAsnArgAsnLysTrpAspAsnArgMet 20
Db      1171 TCAGCTTCGGGACTCGCTCTCTCTATCCAAACCGGAATAAATGGGACATCGTATG 1230

QY      21 SerAlaMetIleProGluIleAsp 28
Db      1231 TCGCGCATGATACCAAGATCGAT 1254

RESULT 4
US-10-326-184-8
; Sequence 8, Application US/10326184
; Publication No. US20030163847A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Company
; TITLE OF INVENTION: REVERSIBLE MALE STERILITY IN TRANSGENIC PLANTS BY EXPRESSION OF CY
; FILE REFERENCE: MTC6781.1
; CURRENT APPLICATION NUMBER: US/10/326,184
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US 60/343,129
; PRIOR FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 1506
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-326-184-8

Alignment Scores:
Pred. No.:      1,2e-15      Length:      1506
Score:          147.00      Matches:      28
Percent Similarity: 100.00%      Conservatives: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              16      Gaps:      0

US-10-014-101B-32 (1-28) x US-10-326-184-8 (1-1506)
```

Qy 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20  
Db 1171 TCAGCTTCGGGACTCGCTCTCTCTATCCAAACCGGATAAATGGGACAATCGGTATG 1230  
Qy 21 SerAlaMetIleProGluIleAsp 28  
Db 1231 TCGCGGATGATACAGAGATCGAT 1254

## RESULT 5

US-10-311-453-26  
; Sequence 26, Application US/10311453  
; Publication No. US20040031073A1  
; GENERAL INFORMATION:  
; APPLICANT: Schmilling, Thomas  
; APPLICANT: Werner, Tom s  
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and  
; TITLE OF INVENTION: physiology  
; FILE REFERENCE: 1226-4  
; CURRENT APPLICATION NUMBER: US/10/311,453  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: EP 00870132.8  
; PRIOR FILING DATE: 2000-06-16  
; PRIOR APPLICATION NUMBER: US 60/258,415  
; PRIOR FILING DATE: 2000-12-27  
; PRIOR APPLICATION NUMBER: EP 01870053.4  
; PRIOR FILING DATE: 2001-03-16  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 26  
; LENGTH: 1506  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-10-311-453-26

## Alignment Scores:

Pred. No.:	1.2e-15	Length:	1506
Score:	147.00	Matches:	28
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	17	Gaps:	0

US-10-014-101B-32 (1-28) x US-10-311-453-26 (1-1506)

Qy 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20  
Db 1171 TCAGCTTCGGGACTCGCTCTCTCTATCCAAACCGGATAAATGGGACAATCGGTATG 1230  
Qy 21 SerAlaMetIleProGluIleAsp 28  
Db 1231 TCGCGGATGATACAGAGATCGAT 1254

## RESULT 6

US-09-938-842A-2089  
; Sequence 2089, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; TITLE OF INVENTION: SAME, AND METHODS OF USE  
; FILE REFERENCE: SCRI1300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22

; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 2089  
; LENGTH: 1575  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-2089

## Alignment Scores:

Pred. No.:	1.07e-08	Length:	1575
Score:	108.00	Matches:	19
Percent Similarity:	78.57%	Conservative:	3
Best Local Similarity:	67.86%	Mismatches:	6
Query Match:	73.47%	Indels:	0
DB:	9	Gaps:	0

US-10-014-101B-32 (1-28) x US-09-938-842A-2089 (1-1575)

Qy 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20  
Db 1240 TCAACTTCGGGTGTACTCTCTCTATCCAAACCGGATAAATGGGACAATCGGTATG 1299  
Qy 21 SerAlaMetIleProGluIleAsp 28  
Db 1300 TCAACGATGACACCGGACGAGAT 1323

## RESULT 7

US-09-938-842A-2089  
; Sequence 2089, Application US/09938842A  
; Publication No. US20040009476A9  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SCRI1300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 2089  
; LENGTH: 1575  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-2089

## Alignment Scores:

Pred. No.:	1.07e-08	Length:	1575
Score:	108.00	Matches:	19
Percent Similarity:	78.57%	Conservative:	3
Best Local Similarity:	67.86%	Mismatches:	6
Query Match:	73.47%	Indels:	0
DB:	11	Gaps:	0

US-10-014-101B-32 (1-28) x US-09-938-842A-2089 (1-1575)

Qy 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20  
Db 1240 TCAACTTCGGGTGTACTCTCTCTATCCAAACCGGATAAATGGGACAATCGGTATG 1299  
Qy 21 SerAlaMetIleProGluIleAsp 28  
Db 1300 TCAACGATGACACCGGACGAGAT 1323

## RESULT 8

US-10-014-101-28  
; Sequence 28, Application US/10014101

```
; Publication No. US20030074698A1
; GENERAL INFORMATION:
; APPLICANT: Schmulling, Thomas
; APPLICANT: Werner, Tom s
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and
; TITLE OF INVENTION: physiology
; FILE REFERENCE: 1195-2
; CURRENT APPLICATION NUMBER: US/10/014,101
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: PCT/EP01/06833
; PRIOR FILING DATE: 2001-06-16
; PRIOR APPLICATION NUMBER: EP 00870132.8
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/258,415
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: EP 01870053.4
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 1575
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-014-101-28

Alignment Scores:
Pred. No.: 1.07e-08 Length: 1575
Score: 108.00 Matches: 19
Percent Similarity: 78.57% Conservative: 3
Best Local Similarity: 67.86% Mismatches: 6
Query Match: 73.47% Indels: 0
DB: 14 Gaps: 0

US-10-014-101B-32 (1-28) x US-10-014-101-28 (1-1575)

Qy 1 SerAlaSerGlyLeuAlaLeuLeuTyrProThrAsnArgAsnLysTrpAspAsnArgMet 20
Db 1240 TCAACTTCTGGTGTTACTCTCTTATCCCAACCGAACAATGGAAACACCGCATG 1299

Qy 21 SerAlaMetIleProGluIleAsp 28
Db 1300 TCAACGATGACACCGGACGAAGAT 1323

RESULT 9
US-10-326-184-10
; Sequence 10, Application US/10326184
; Publication No. US20030163847A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Company
; TITLE OF INVENTION: REVERSIBLE MALE STERILITY IN TRANSGENIC PLANTS BY EXPRESSION OF
; TITLE OF INVENTION: OXIDASE 1
; FILE REFERENCE: MTC6781.1
; CURRENT APPLICATION NUMBER: US/10/326,184
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US 60/343,129
; PRIOR FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 1575
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-326-184-10

Alignment Scores:
Pred. No.: 1.07e-08 Length: 1575
Score: 108.00 Matches: 19
Percent Similarity: 78.57% Conservative: 3
Best Local Similarity: 67.86% Mismatches: 6
Query Match: 73.47% Indels: 0
DB: 16 Gaps: 0

US-10-014-101B-32 (1-28) x US-10-326-184-10 (1-1575)
```

```
; Publication No. US20030074698A1
; GENERAL INFORMATION:
; APPLICANT: Schmulling, Thomas
; APPLICANT: Werner, Tom s
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and
; TITLE OF INVENTION: physiology
; FILE REFERENCE: 1195-2
; CURRENT APPLICATION NUMBER: US/10/014,101
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: PCT/EP01/06833
; PRIOR FILING DATE: 2001-06-16
; PRIOR APPLICATION NUMBER: EP 00870132.8
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/258,415
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: EP 01870053.4
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 1575
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-014-101-28

Alignment Scores:
Pred. No.: 1.07e-08 Length: 1575
Score: 108.00 Matches: 19
Percent Similarity: 78.57% Conservative: 3
Best Local Similarity: 67.86% Mismatches: 6
Query Match: 73.47% Indels: 0
DB: 14 Gaps: 0

US-10-014-101B-32 (1-28) x US-10-311-453-28 (1-1575)

Qy 1 SerAlaSerGlyLeuAlaLeuLeuTyrProThrAsnArgAsnLysTrpAspAsnArgMet 20
Db 1240 TCAACTTCTGGTGTTACTCTCTTATCCCAACCGAACAATGGAAACACCGCATG 1299

Qy 21 SerAlaMetIleProGluIleAsp 28
Db 1300 TCAACGATGACACCGGACGAAGAT 1323

RESULT 11
US-10-326-184-20
; Sequence 20, Application US/10326184
; Publication No. US20030163847A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Company
; TITLE OF INVENTION: REVERSIBLE MALE STERILITY IN TRANSGENIC PLANTS BY EXPRESSION OF C
; TITLE OF INVENTION: OXIDASE 1
; FILE REFERENCE: MTC6781.1
; CURRENT APPLICATION NUMBER: US/10/326,184
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US 60/343,129
; PRIOR FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 1655
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-326-184-20
```

```
Alignment Scores:
Pred. No.: 1.14e-08 Length: 1655
Score: 108.00 Matches: 19
Percent Similarity: 78.57% Conservative: 3
Best Local Similarity: 67.86% Mismatches: 6
Query Match: 73.47% Indels: 0
DB: 16 Gaps: 0

US-10-014-101B-32 (1-28) x US-10-326-184-20 (1-1655)
QY 1 SerAlaSerGlyLeuAlaLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20
    |||||
DB 1240 TCACCTTCGGTGTACTCTCTCTATCCCAACCGAACAATGGAACACCGCATG 1299
    |||||

QY 21 SerAlaMetIleProGluIleAsp 28
    |||||
DB 1300 TCAACGATGACACCGGACGAAGAT 1323
    |||||

RESULT 12
US-10-014-101-3
; Sequence 3, Application US/10014101
; Publication No. US20030074698A1
; GENERAL INFORMATION:
; APPLICANT: Schmulling, Thomas
; APPLICANT: Werner, Tom s
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and
; FILE REFERENCE: 1195-2
; CURRENT APPLICATION NUMBER: US/10/014,101
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: PCT/EP01/06833
; PRIOR FILING DATE: 2001-06-16
; PRIOR APPLICATION NUMBER: EP 00870132.8
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/258,415
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: EP 01870053.4
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2991
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-014-101-3
Alignment Scores:
Pred. No.: 2.67e-06 Length: 2991
Score: 96.50 Matches: 27
Percent Similarity: 48.21% Conservative: 0
Best Local Similarity: 48.21% Mismatches: 1
Query Match: 65.65% Indels: 29
DB: 17 Gaps: 1

US-10-014-101B-32 (1-28) x US-10-311-453-3 (1-2991)
QY 1 SerAlaSerGlyLeuAlaLeuTyProThrAsnArgAsnLys----- 15
    |||||
DB 2573 TCAGCTTCGGGACTCGCTCTCTCTATCCCAACCGGAATAA-GTACATACCTTCTCT 2631
    |||||

QY 15 ----- 15
DB 2632 CATTCATATTTATCTTCAAGAACCAAGTAATAATTTCTATGAACGTATTATGCTGTT 2691
    |||||

QY 16 -----TrpAspAsnArgMetSerAlaMetIleProGluIleAsp 28
    |||||
DB 2692 ATTGTTAGTGGGACAAATCGTATGTCGGCATGATACCAAGATCGAT 2739
    |||||

RESULT 14
US-10-014-101-27
; Sequence 27, Application US/10014101
; Publication No. US20030074698A1
; GENERAL INFORMATION:
; APPLICANT: Schmulling, Thomas
; APPLICANT: Werner, Tom s
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and
; FILE REFERENCE: 1195-2
; CURRENT APPLICATION NUMBER: US/10/014,101
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: PCT/EP01/06833
; PRIOR FILING DATE: 2001-06-16
; PRIOR APPLICATION NUMBER: EP 00870132.8
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/258,415
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: EP 01870053.4
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 1572
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-014-101-27
Alignment Scores:
Pred. No.: 2.67e-06 Length: 2991
Score: 96.50 Matches: 27
Percent Similarity: 48.21% Conservative: 0
Best Local Similarity: 48.21% Mismatches: 1
Query Match: 65.65% Indels: 29
DB: 14 Gaps: 1

US-10-014-101B-32 (1-28) x US-10-014-101-3 (1-2991)
QY 1 SerAlaSerGlyLeuAlaLeuTyProThrAsnArgAsnLys----- 15
    |||||
DB 2573 TCAGCTTCGGGACTCGCTCTCTCTATCCCAACCGGAATAA-GTACATACCTTCTCT 2631
    |||||

QY 15 ----- 15
DB 2632 CATTCATATTTATCTTCAAGAACCAAGTAATAATTTCTATGAACGTATTATGCTGTT 2691
    |||||

QY 16 -----TrpAspAsnArgMetSerAlaMetIleProGluIleAsp 28
    |||||
DB 2692 ATTGTTAGTGGGACAAATCGTATGTCGGCATGATACCAAGATCGAT 2739
    |||||

RESULT 13
US-10-311-453-3
; Sequence 3, Application US/10311453
; Publication No. US20040031073A1
```



Publication No. US20040031072A1

GENERAL INFORMATION:

APPLICANT: La Rosa Thomas J.

APPLICANT: Kovalic David K.

APPLICANT: Zhou Yihua

APPLICANT: Cao Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53223)B

CURRENT APPLICATION NUMBER: US/10/424,599

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 285684

SEQ ID NO 1725

LENGTH: 1142

TYPE: DNA

ORGANISM: Glycine max

FEATURE:

OTHER INFORMATION: Clone ID: PAT\_MRT3847\_101557C.1

US-10-424-599-1725

Alignment Scores:

Pred. No.: 7,46e-06 Length: 1142

Score: 91.00 Matches: 15

Percent Similarity: 78.57% Conservative: 7

Best Local Similarity: 53.57% Mismatches: 6

Query Match: 61.90% Indels: 0

DB: 17 Gaps: 0

US-10-014-101B-32 (1-28) x US-10-424-599-1725 (1-1142)

QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20

Db 611 ACTAGCAACGGCCGCTCTTATCTACCCAGTAACAAATCAAAATGGGACACAGAACT 670

QY 21 SerAlaMetIleProGluIleAsp 28

Db 671 TCTGTGTGATTCACAGGAGAT 594

RESULT 19

US-10-425-114-8021

Sequence 8021, Application US/10425114

Publication No. US20040034888A1

GENERAL INFORMATION:

APPLICANT: Liu, Jingdong

APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.

APPLICANT: Screen, Steven E

APPLICANT: Tabaska, Jack E

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53313)B

CURRENT APPLICATION NUMBER: US/10/425,114

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 73128

SEQ ID NO 8021

LENGTH: 1279

TYPE: DNA

ORGANISM: Glycine max

FEATURE:

OTHER INFORMATION: Clone ID: 700731265\_FLI

US-10-425-114-8021

Alignment Scores:

Pred. No.: 8,61e-06 Length: 1279

Score: 91.00 Matches: 15

Percent Similarity: 78.57% Conservative: 7

Best Local Similarity: 53.57% Mismatches: 6

Query Match: 61.90% Indels: 0

DB: 17 Gaps: 0

US-10-014-101B-32 (1-28) x US-10-425-114-8021 (1-1279)

QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20

Db 689 ACAAGCAACGGCCGCTCTTATCTACCCAGTAACAAATCAAAATGGGACACAGAACT 748

QY 21 SerAlaMetIleProGluIleAsp 28

Db 749 TCTGTGTGATTCACAGGAGAT 772

RESULT 20

US-10-424-599-1724

Sequence 1724, Application US/10424599

Publication No. US20040031072A1

GENERAL INFORMATION:

APPLICANT: La Rosa Thomas J

APPLICANT: Kovalic David K

APPLICANT: Zhou Yihua

APPLICANT: Cao Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53223)B

CURRENT APPLICATION NUMBER: US/10/424,599

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 285684

SEQ ID NO 1724

LENGTH: 1587

TYPE: DNA

ORGANISM: Glycine max

FEATURE:

NAME/KEY: unsure

LOCATION: (1)..(1587)

OTHER INFORMATION: unsure at all n locations

FEATURE:

OTHER INFORMATION: Clone ID: PAT\_MRT3847\_101556C.1

US-10-424-599-1724

Alignment Scores:

Pred. No.: 1,13e-05 Length: 1587

Score: 91.00 Matches: 15

Percent Similarity: 78.57% Conservative: 7

Best Local Similarity: 53.57% Mismatches: 6

Query Match: 61.90% Indels: 0

DB: 17 Gaps: 0

US-10-014-101B-32 (1-28) x US-10-424-599-1724 (1-1587)

QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20

Db 726 ACAAGCAACGGCCGCTCTTATCTACCCAGTAACAAATCAAAATGGGACACAGAACT 785

QY 21 SerAlaMetIleProGluIleAsp 28

Db 786 TCTGTGTGATTCACAGGAGAT 809

RESULT 21

US-10-437-963-46009

Sequence 46009, Application US/10437963

Publication No. US20040123343A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

APPLICANT: Wu, Wei

APPLICANT: Boukharov, Andrey A.

APPLICANT: Barbazuk, Brad

APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53221)B

CURRENT APPLICATION NUMBER: US/10/437,963

CURRENT FILING DATE: 2003-05-14

NUMBER OF SEQ ID NOS: 204966

SEQ ID NO 46009

```
; LENGTH: 2054
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_48918C.1
US-10-437-963-46009

Alignment Scores:
Pred. No.:          5.35e-05      Length:      2054
Score:              88.00         Matches:      16
Percent Similarity: 73.08%        Conservative: 3
Best Local Similarity: 61.54%      Mismatches:  7
Query Match:        59.86%        Indels:      0
DB:                 18           Gaps:        0

US-10-014-101B-32 (1-28) x US-10-437-963-46009 (1-2054)

Qy 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20
Db 1210 AGCACAGTGGTCCCATCTACTGCTTTACCCAGTGAACAGAACCAAGTGGGACACAGAAC 1269

Qy 21 SerAlaMetIleProGlu 26
Db 1270 TCAGTGGTCATACCATGAT 1287

RESULT 22
US-10-437-963-43981
; Sequence 43981, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 43981
; LENGTH: 1677
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_47086C.1
US-10-437-963-43981

Alignment Scores:
Pred. No.:          6.23e-05      Length:      1677
Score:              87.00         Matches:      16
Percent Similarity: 74.07%        Conservative: 4
Best Local Similarity: 59.26%      Mismatches:  7
Query Match:        59.18%        Indels:      0
DB:                 18           Gaps:        0

US-10-014-101B-32 (1-28) x US-10-437-963-43981 (1-1677)

Qy 2 AlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSer 21
Db 1318 GCCATGGGCCCGCTCTCATCTACCCATGAACCGCAACGACGACAGTAAACATGTGCG 1377

Qy 22 AlaMetIleProGluIleAsp 28
Db 1378 GCAGTGTATCCCGACCGACGAC 1398

RESULT 23
US-10-817-483-28
; Sequence 28, Application US/10817483

; Publication No. US20040237147A1
; GENERAL INFORMATION:
; APPLICANT: Habben, Jeffrey E.
; APPLICANT: Zinselmeyer, Christopher
; APPLICANT: Tomes, Dwight
; APPLICANT: Abbott, Shane
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Niu, Xiaomu
; TITLE OF INVENTION: Modulation of Cytokinin Activity in
; FILE REFERENCE: 0803R
; CURRENT APPLICATION NUMBER: US/10/817,483
; CURRENT FILING DATE: 2004-04-02
; PRIOR APPLICATION NUMBER: US 60/460,718
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US 09/545,334
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: US 60/129,844
; PRIOR FILING DATE: 1999-04-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 1617
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)....(1617)
US-10-817-483-28

Alignment Scores:
Pred. No.:          8.96e-05      Length:      1617
Score:              86.00         Matches:      13
Percent Similarity: 80.77%        Conservative: 8
Best Local Similarity: 50.00%      Mismatches:  5
Query Match:        58.50%        Indels:      0
DB:                 18           Gaps:        0

US-10-014-101B-32 (1-28) x US-10-817-483-28 (1-1617)

Qy 3 SerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22
Db 1246 GCTGGGATCATCTCATGTACCCATGAACAAAGGATAGTGGGACGCCGATGACAGCG 1305

Qy 23 MetIleProGluIleAsp 28
Db 1306 ATGACCCCGACCGACGAC 1323

RESULT 24
US-10-425-114-10804
; Sequence 10804, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 10804
; LENGTH: 1023
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700978152_FLI
US-10-425-114-10804
```

```

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 83827
; LENGTH: 1673
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_83121C.1
US-10-437-963-83827

Alignment Scores:
Pred. No.: 0.000141 Length: 1673
Score: 85.00 Matches: 14
Percent Similarity: 80.00% Conservative: 6
Best Local Similarity: 56.00% Mismatches: 5
Query Match: 57.82% Indels: 0
DB: 18 Gaps: 0

US-10-014-101B-32 (1-28) x US-10-437-963-83827 (1-1673)

Qy 4 GlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSerAlaMet 23
Db 1341 GGCCTCATCTCTATACCCCTGAATAAGGACATGTGGGATGACAGGATGACGGCGATG 1400
Qy 24 IleProGluIleAsp 28
Db 1401 ACGCCGGACGAGGAC 1415

RESULT 27
US-10-425-114-15562
; Sequence 15562, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 15562
; LENGTH: 1042
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3060-025-Fl1_FLI
US-10-425-114-15562

Alignment Scores:
Pred. No.: 0.000175 Length: 1042
Score: 83.00 Matches: 14
Percent Similarity: 76.92% Conservative: 6
Best Local Similarity: 53.85% Mismatches: 6
Query Match: 56.46% Indels: 0
DB: 17 Gaps: 0

US-10-014-101B-32 (1-28) x US-10-425-114-15562 (1-1042)

Qy 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20
Db 156 AGCAACAATGGTCCTATATTTGCTTATCCAGTGACAAATCAAGTGGGACACAAACG 215
Qy 21 SerAlaMetIleProGlu 26
Db 216 TCAGTGGTCATACCAT 233

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 83827
; LENGTH: 1673
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_83121C.1
US-10-437-963-83827

Alignment Scores:
Pred. No.: 0.000141 Length: 1673
Score: 85.00 Matches: 14
Percent Similarity: 80.00% Conservative: 6
Best Local Similarity: 56.00% Mismatches: 5
Query Match: 57.82% Indels: 0
DB: 18 Gaps: 0

US-10-014-101B-32 (1-28) x US-10-425-114-10804 (1-1023)

Qy 3 SerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22
Db 514 AGTGGGCTATCTCATATACCTATGACAAACAAAGTGGGACCATCGAAGCTCAGT 573

RESULT 25
US-10-424-599-19583
; Sequence 19583, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 19583
; LENGTH: 1045
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_117689C.1
US-10-424-599-19583

Alignment Scores:
Pred. No.: 7.76e-05 Length: 1045
Score: 85.00 Matches: 15
Percent Similarity: 73.08% Conservative: 4
Best Local Similarity: 57.69% Mismatches: 7
Query Match: 57.82% Indels: 0
DB: 17 Gaps: 0

US-10-014-101B-32 (1-28) x US-10-424-599-19583 (1-1045)

Qy 3 SerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22
Db 536 AGTGGGCTATCTCATATACCTATGACAAACAAAGTGGGACCATCGAAGCTCAGT 595

RESULT 26
US-10-437-963-83827
; Sequence 83827, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

```



```
; TITLE OF INVENTION: physiology
; FILE REFERENCE: 1195-2
; CURRENT APPLICATION NUMBER: US/10/014,101
; CURRENT FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: PCT/EP01/06833
; PRIOR FILING DATE: 2001-06-16
; PRIOR APPLICATION NUMBER: EP 00870132.8
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/258,415
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: EP 01870053.4
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 2782
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-014-101-7

Alignment Scores:
Pred. No.: 0.000608      Length: 2782
Score: 83.00             Matches: 19
Percent Similarity: 37.93% Conservative: 3
Best Local Similarity: 32.76% Mismatches: 6
Query Match: 56.46%      Indels: 30
DB: 14                   Gaps: 1

US-10-014-101B-32 (1-28) x US-10-014-101-7 (1-2782)
QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLys-----15
Db 2357 TCAACTTCTGGTGTACTCTCTCTATCCCAACCAACGAAACAAAGTAAATATTACTTTT 2416
QY 15 -----15
Db 2417 TGATTTTGTATTGTTTAAAGTATATCCCAATAATGTTAAATTTGTTAAACAAGATT 2476
QY 16 -----TrpAspAsnArgMetSerAlaMetIleProGluLeuAep 28
Db 2477 TATTTTAAATAGATGGAACACCGCATGTCAACGATGACACCGGACGAAGAT 2530

RESULT 33
US-10-424-599-62631
; Sequence 62631, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 62631
; LENGTH: 816
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(816)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_27570C.1
US-10-424-599-62631

Alignment Scores:
Pred. No.: 0.000193      Length: 816
Score: 82.00             Matches: 15
Percent Similarity: 78.26% Conservative: 3
Best Local Similarity: 65.22% Mismatches: 5
Query Match: 55.78%      Indels: 0
DB: 17                   Gaps: 0

US-10-014-101B-32 (1-28) x US-10-424-599-62631 (1-816)
QY 4 GlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSerAlaMet 23
Db 90 GGACCACTCTGGTTTATCCCATGATAGAAAACAAAGTGTACATAGGATGTCAGCATCT 149
QY 24 IleProGlu 26
Db 150 ATACCAAGAC 158

RESULT 34
US-10-437-963-38170
; Sequence 38170, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
```

```
; TITLE OF INVENTION: physiology
; FILE REFERENCE: 1195-2
; CURRENT APPLICATION NUMBER: US/10/014,101
; CURRENT FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: PCT/EP01/06833
; PRIOR FILING DATE: 2001-06-16
; PRIOR APPLICATION NUMBER: EP 00870132.8
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/258,415
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: EP 01870053.4
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 2782
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-014-101-7

Alignment Scores:
Pred. No.: 0.000608      Length: 2782
Score: 83.00             Matches: 19
Percent Similarity: 37.93% Conservative: 3
Best Local Similarity: 32.76% Mismatches: 6
Query Match: 56.46%      Indels: 30
DB: 14                   Gaps: 1

US-10-014-101B-32 (1-28) x US-10-014-101-7 (1-2782)
QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLys-----15
Db 2357 TCAACTTCTGGTGTACTCTCTCTATCCCAACCAACGAAACAAAGTAAATATTACTTTT 2416
QY 15 -----15
Db 2417 TGATTTTGTATTGTTTAAAGTATATCCCAATAATGTTAAATTTGTTAAACAAGATT 2476
QY 16 -----TrpAspAsnArgMetSerAlaMetIleProGluLeuAep 28
Db 2477 TATTTTAAATAGATGGAACACCGCATGTCAACGATGACACCGGACGAAGAT 2530

RESULT 32
US-10-311-453-7
; Sequence 7, Application US/10311453
; Publication No. US20040031073A1
; GENERAL INFORMATION:
; APPLICANT: Schilling, Thomas
; APPLICANT: Werner, Tom B
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and
; FILE OF INVENTION: physiology
; FILE REFERENCE: 1226-4
; CURRENT APPLICATION NUMBER: US/10/311,453
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: EP 00870132.8
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/258,415
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: EP 01870053.4
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 2782
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-311-453-7

Alignment Scores:
Pred. No.: 0.000608      Length: 2782
Score: 83.00             Matches: 19
Percent Similarity: 37.93% Conservative: 3
Best Local Similarity: 32.76% Mismatches: 6
```

APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 38170  
LENGTH: 1605  
TYPE: DNA  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_418332C.1  
US-10-437-963-38170

Alignment Scores:  
Pred. No.: 0.000686 Length: 1605  
Score: 81.00 Matches: 15  
Percent Similarity: 67.86% Conservative: 4  
Best Local Similarity: 53.57% Mismatches: 9  
Query Match: 55.10% Indels: 0  
DB: 18 Gaps: 0

US-10-014-101B-32 (1-28) x US-10-437-963-38170 (1-1605)

Qy 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20  
Db 1231 ACCGCGGTGGCCCACTCTCTACCCAGTGAACAAATCAAAGTGGGACATCAAACA 1209  
Qy 21 SerAlaMetIleProGlu 28  
Db 1291 TCGTGTGTGACGCGGAGGAGGAC 1314

## RESULT 35

US-10-014-101-30  
Sequence 30, Application US/10014101  
Publication No. US20030074698A1  
GENERAL INFORMATION:  
APPLICANT: Schmulling, Thomas  
APPLICANT: Werner, Tom s  
TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and  
FILE REFERENCE: 1195-2  
CURRENT APPLICATION NUMBER: US/10/014,101  
CURRENT FILING DATE: 2001-12-10  
PRIOR APPLICATION NUMBER: PCT/EP01/06833  
PRIOR FILING DATE: 2001-06-16  
PRIOR APPLICATION NUMBER: EP 00870132.8  
PRIOR FILING DATE: 2000-06-16  
PRIOR APPLICATION NUMBER: US 60/258,415  
PRIOR FILING DATE: 2000-12-27  
PRIOR APPLICATION NUMBER: EP 01870053.4  
PRIOR FILING DATE: 2001-03-16  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 30  
LENGTH: 1515  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-10-014-101-30

Alignment Scores:  
Pred. No.: 0.000959 Length: 1515  
Score: 80.00 Matches: 12  
Percent Similarity: 80.77% Conservative: 9  
Best Local Similarity: 46.15% Mismatches: 5  
Query Match: 54.42% Indels: 0  
DB: 14 Gaps: 0

US-10-014-101B-32 (1-28) x US-10-014-101-30 (1-1515)

Qy 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20

Db 1150 ACAAGCAACGCCCGAGTCATCGTCTACCCAGTGAACAAATCAAAGTGGGACATCAAACA 1209  
Qy 21 SerAlaMetIleProGlu 26  
Db 1210 TCAGCAGTAACACCGGAG 1227  
RESULT 36  
US-10-311-453-30  
Sequence 30, Application US/10311453  
Publication No. US20040031073A1  
GENERAL INFORMATION:  
APPLICANT: Schmulling, Thomas  
APPLICANT: Werner, Tom s  
TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and  
FILE REFERENCE: 1226-4  
CURRENT APPLICATION NUMBER: US/10/311,453  
CURRENT FILING DATE: 2002-12-16  
PRIOR APPLICATION NUMBER: EP 00870132.8  
PRIOR FILING DATE: 2000-06-16  
PRIOR APPLICATION NUMBER: US 60/258,415  
PRIOR FILING DATE: 2000-12-27  
PRIOR APPLICATION NUMBER: EP 01870053.4  
PRIOR FILING DATE: 2001-03-16  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 30  
LENGTH: 1515  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-10-311-453-30

## Alignment Scores:

Pred. No.: 0.000959 Length: 1515  
Score: 80.00 Matches: 12  
Percent Similarity: 80.77% Conservative: 9  
Best Local Similarity: 46.15% Mismatches: 5  
Query Match: 54.42% Indels: 0  
DB: 17 Gaps: 0

US-10-014-101B-32 (1-28) x US-10-311-453-30 (1-1515)

Qy 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20  
Db 1150 ACAAGCAACGCCCGAGTCATCGTCTACCCAGTGAACAAATCAAAGTGGGACATCAAACA 1209  
Qy 21 SerAlaMetIleProGlu 26  
Db 1210 TCAGCAGTAACACCGGAG 1227

## RESULT 37

US-10-021-323-9662  
Sequence 9662, Application US/10021323  
Publication No. US20040123340A1  
GENERAL INFORMATION:  
APPLICANT: Deikman, Jill  
APPLICANT: Feng, Paul C.C.  
APPLICANT: Fincher, Karen L.  
APPLICANT: Ziegler, Todd E.  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(52274)B  
CURRENT APPLICATION NUMBER: US/10/021,323  
CURRENT FILING DATE: 2001-12-12  
PRIOR APPLICATION NUMBER: US 60/255, 619  
PRIOR FILING DATE: 2000-12-14  
NUMBER OF SEQ ID NOS: 17880  
SEQ ID NO 9662  
LENGTH: 482  
TYPE: DNA  
ORGANISM: Gossypium hirsutum  
FEATURE:

OTHER INFORMATION: Clone ID: LIB3828-026-Q6-K6-E7  
US-10-021-323-9662

Alignment Scores:  
Pred. No.: 0.000338 Length: 482  
Score: 79.00 Matches: 13  
Percent Similarity: 75.00% Conservative: 5  
Best Local Similarity: 54.17% Mismatches: 6  
Query Match: 53.74% Indels: 0  
DB: 18 Gaps: 0

US-10-014-101B-32 (1-28) x US-10-021-323-9662 (1-482)

Qy 3 SerGlyLeuAlaLeuTyrProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22

Db 107 AGTGGCCCAATTCTCTATCAATGCAACAAGCAAAATGGGACCATAGGAGCTCCGTG 166

Qy 23 MetIleProGlu 26

Db 167 GTGACACCGGAT 178

RESULT 38

US-10-021-323-10415/c  
Sequence 10415, Application US/10021323

Publication No. US20040123340A1

GENERAL INFORMATION:

APPLICANT: Deikman, Jill

APPLICANT: Feng, Paul C. C.

APPLICANT: Fincher, Karen L.

APPLICANT: Ziegler, Todd E.

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with

TITLE OF INVENTION: Plants

FILE REFERENCE: 38-21(52274)B

CURRENT APPLICATION NUMBER: US/10/021.323

PRIOR FILING DATE: 2001-12-12

PRIOR APPLICATION NUMBER: US 60/255, 619

PRIOR FILING DATE: 2000-12-14

NUMBER OF SEQ ID NOS: 17880

SEQ ID NO 10415

LENGTH: 484

TYPE: DNA

ORGANISM: Gossypium hirsutum

FEATURE:

NAME/KEY: unsure

LOCATION: (1)..(484)

OTHER INFORMATION: unsure at all n locations

OTHER INFORMATION: Clone ID: LIB3828-026-Q6-N6-E7

US-10-021-323-10415

Alignment Scores:

Pred. No.: 0.00034 Length: 484

Score: 79.00 Matches: 13

Percent Similarity: 75.00% Conservative: 5

Best Local Similarity: 54.17% Mismatches: 6

Query Match: 53.74% Indels: 0

DB: 18 Gaps: 0

US-10-014-101B-32 (1-28) x US-10-021-323-10415 (1-484)

Qy 3 SerGlyLeuAlaLeuTyrProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22

Db 376 AGTGGCCCAATTCTCTATCAATGCAACAAGCAAAATGGGACCATAGGAGCTCCGTG 317

Qy 23 MetIleProGlu 26

Db 316 GTGACACCGGAT 305

RESULT 39

US-10-767-795-2077

Sequence 2077, Application US/10767795

Publication No. US20040181830A1

GENERAL INFORMATION:

APPLICANT: Kovalic, David K.

APPLICANT: Cao, Yongwei  
APPLICANT: Zhou, Yihua  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement  
FILE REFERENCE: 38-21(53534)B  
CURRENT APPLICATION NUMBER: US/10/767,795  
CURRENT FILING DATE: 2004-01-30  
NUMBER OF SEQ ID NOS: 117596  
SEQ ID NO 2077  
LENGTH: 788  
TYPE: DNA  
ORGANISM: Gossypium hirsutum  
FEATURE:

OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C36643\_1

US-10-767-795-2077

Alignment Scores:

Pred. No.: 0.000631 Length: 788

Score: 79.00 Matches: 13

Percent Similarity: 75.00% Conservative: 5

Best Local Similarity: 54.17% Mismatches: 6

Query Match: 53.74% Indels: 0

DB: 18 Gaps: 0

US-10-014-101B-32 (1-28) x US-10-767-795-2077 (1-788)

Qy 3 SerGlyLeuAlaLeuTyrProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22

Db 412 AGTGGCCCAATTCTCTATCAATGCAACAAGCAAAATGGGACCATAGGAGCTCCGTG 471

Qy 23 MetIleProGlu 26

Db 472 GTGACACCGGAT 483

RESULT 40

US-10-014-101-29

Sequence 29, Application US/10014101

Publication No. US20030074698A1

GENERAL INFORMATION:

APPLICANT: Schumling, Thomas

APPLICANT: Werner, Tom S

TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and

TITLE OF INVENTION: physiology

FILE REFERENCE: 1195-2

CURRENT APPLICATION NUMBER: US/10/014,101

CURRENT FILING DATE: 2001-12-10

PRIOR APPLICATION NUMBER: PCT/EP01/06833

PRIOR FILING DATE: 2001-06-16

PRIOR APPLICATION NUMBER: EP 00870132.8

PRIOR FILING DATE: 2000-06-16

PRIOR APPLICATION NUMBER: US 60/258,415

PRIOR FILING DATE: 2000-12-27

PRIOR APPLICATION NUMBER: EP 01870053.4

PRIOR FILING DATE: 2001-03-16

NUMBER OF SEQ ID NOS: 36

SOFTWARE: Patent in Ver. 2.1

SEQ ID NO 29

LENGTH: 1611

TYPE: DNA

ORGANISM: Arabidopsis thaliana

US-10-014-101-29

Alignment Scores:

Pred. No.: 0.00354 Length: 1611

Score: 77.00 Matches: 13

Percent Similarity: 75.00% Conservative: 5

Best Local Similarity: 54.17% Mismatches: 6

Query Match: 52.38% Indels: 0

DB: 14 Gaps: 0

US-10-014-101B-32 (1-28) x US-10-014-101-29 (1-1611)

Qy 3 SerGlyLeuAlaLeuTyrProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22

Db 1213 AGTGGCCCTATTCTTATCTACCCCATGAACAAAGACAAATGGGACGAGGAGCTCAGCC 1272  
Qy 23 MetIleProGlu 26  
Db 1273 GTGACGCCGGAT 1284

Search completed: February 18, 2005, 14:57:53  
Job time : 295 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 18, 2005, 09:14:12 ; Search time 7200 Seconds  
(without alignments)  
148.028 Million cell updates/sec

Title: US-10-014-101b-32

Perfect score: 147

Sequence: 1 SASGLALLYPTNRKNKWNMSAMPEID 28

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Zgapop 6.0 , Zgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgm2\_1/USPTO.spool/US10014101/runat\_16022005\_075934\_8314/app\_query.fasta\_1.199  
-DB=EST -OPMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10014101 @CGN 1 1 2607 @runat\_16022005\_075934\_8314 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
C 1	102	69.4	418	5	BP576449
C 2	100	68.0	593	7	CO979507
C 3	96	65.3	1594	3	CNSOAIIGX
C 4	93	63.3	806	7	CF393540
C 5	92	62.6	539	4	BM138354
C 6	91	61.9	697	7	CF398227
C 7	88	59.9	401	1	AV810994
C 8	88	59.9	505	4	BI927011
C 9	88	59.9	642	4	BI926735

C 10	88	59.9	757	7	COL19329
C 11	88	59.9	795	7	COL17913
C 12	88	59.9	807	6	CA766652
C 13	88	59.9	824	7	COL121610
C 14	88	59.9	1578	9	CL960490
C 15	87	59.2	772	6	CB620527
C 16	87	59.2	863	7	CO435706
C 17	87	59.2	1671	9	CL957466
C 18	86.5	58.8	436	1	AU227876
C 19	86	58.5	438	4	BG518327
C 20	86	58.5	642	4	BI203744
C 21	86	58.5	715	7	BI204708
C 22	86	58.5	826	8	BZ644495
C 23	86	58.5	906	9	CG289642
C 24	86	58.5	963	9	CG351153
C 25	85	57.8	584	6	CB913923
C 26	85	57.8	741	7	CF210007
C 27	85	57.8	773	4	BI970438
C 28	85	57.8	809	7	CF209921
C 29	85	57.8	1587	9	CL964867
C 30	84	57.1	812	7	COL17423
C 31	84	57.1	812	7	COL174338
C 32	83	56.5	481	1	AV536711
C 33	83	56.5	1566	9	CL947867
C 34	82	55.8	575	2	BF520835
C 35	82	55.8	610	9	CC694028
C 36	82	55.8	636	6	CA918993
C 37	82	55.8	753	9	CC693605
C 38	82	55.8	788	6	CF068302
C 39	82	55.8	831	7	CO483046
C 40	82	55.8	856	7	CO485966
C 41	82	55.8	876	9	CG445712
C 42	82	55.8	969	9	CG289653
C 43	81	55.1	468	1	AI774379
C 44	80	54.4	589	7	CO490541
C 45	80	54.4	831	7	CN201821

#### ALIGNMENTS

RESULT 1  
BP576449/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

FEATURES

source

BP576449 RAPL14 Arabidopsis thaliana CDNA clone RAF14-92-C01 3',  
mRNA sequence.

BP576449.1 GI:48992215

EST.

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi

1 (bases 1 to 418)

Seki,M., Narusaka,M., Kamiya,A., Ishida,J., Satou,M., Sakurai,T.,

NakaJima,M., Enju,A., Akiyama,K., Oono,Y., Muramatsu,M.,

Hayashiaki,Y., Kawai,J., Carninci,P., Itoh,M., Ishii,Y.,

Arakawa,T., Shibata,K., Shinagawa,A. and Shinozaki,K.

Functional annotation of a full-length Arabidopsis cDNA collection

Science 296 (5565), 141-145 (2002)

21932900

11910074

Contact: Motoaki Seki

RIKEN Genomic Sciences Center

2-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

Tel: 81-298-36-4359

Fax: 81-298-36-9060

Email: mseki@tc.riken.go.jp

reversed clone; please visit

(http://pfweb.gsc.riken.go.jp/) for further details.

Location/Qualifiers

1. .418

source



REFERENCE  
AUTHORS

1 (bases 1 to 1594)  
Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.  
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:  
A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation

JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Unpublished  
Genoscope.  
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
Direct Submission  
Genoscope.  
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.  
URGV INRA : Clepet C., Caboche M.  
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap  
http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full length  
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES  
source

1. 1594  
Location/Qualifiers  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/strain="Col-0"  
/db\_xref="taxon:3702"  
/clone="GSLTPB57ZE08"  
/tissue type="Flowers and buds"  
/plasmid="pCMVSPORT 6"  
complement (1..1594)  
/gene="At5g56970"

gene

ORIGIN

Alignment Scores:  
Pred. No.: 0.00019 Length: 1594  
Score: 96.00 Matches: 18  
Percent Similarity: 80.77% Conservative: 3  
Best Local Similarity: 69.23% Mismatches: 5  
Query Match: 65.31% Indels: 0  
DB: 3 Gaps: 0

US-10-014-101B-32 (1-28) x CNS0A1GX (1-1594)

QY 3 SerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22  
|||||  
DB 1254 AGCGGCTCTGTTCTGTTATCTATGATCGCAAGTGAATGATCGATGCTGTC 1313  
|||||

QY 23 MetIleProGluIleAep 28  
|||||

DB 1314 GCTATACCCGAGGAAGAT 1331  
|||||

RESULT 4  
CF393540  
LOCUS  
DEFINITION  
RTRD3\_13\_E04\_g1\_A022 Loblolly pine roots recovering from drought  
DR3 Pinus taeda cDNA clone RTRD3\_13\_E04\_A022 5', mRNA sequence.  
ACCESSION  
CF393540  
VERSION  
CF393540.1 GI:34351957  
KEYWORDS  
EST.  
SOURCE  
Pinus taeda (loblolly pine)  
ORGANISM  
Pinus taeda  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.  
1 (bases 1 to 806)

REFERENCE  
AUTHORS

Pratt, L., Cordonnier-Pratt, M.-M., Lorenz, W.W., Dean, J., Gebremedhin, M., Dervinis, C., Martin, T., White, T., Davis, J. and

Neale, D.  
An EST database from loblolly pine (Pinus taeda) roots recovering from drought stress  
Unpublished (2003)  
Other ESTs: RTRD3\_13\_E04\_b1\_A022  
Contact: Cordonnier-Pratt MW  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@uga.edu  
RNA prepared and library constructed by W. Walter Lorenz, School of Forestry, University of Georgia; plant material prepared at the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.  
Seq primer: JENREV (CAGGAACAGCTATGACC).

FEATURES  
Location/Qualifiers  
1. 806  
/organism="Pinus taeda"  
/mol\_type="mRNA"  
/strain="CCLONES"  
/db\_xref="taxon:3352"  
/clone="RTRD3\_13\_E04\_A022"  
/lab host="DH10B-T1 phage-resistant E. coli"  
/clone\_lib="Loblolly pine roots recovering from drought DR3"  
/notes="Vector: pSL1180; Site 1: EcoRI; Site 2: XhoI; The library was prepared from polyA+ RNA from loblolly pine (Pinus taeda) roots recovering from drought. Water was withheld from tamed clones until predawn needle water potential reached -1.75 MPa. Plants were well watered on day 7 and allowed to recover for 2 days, at which time roots were harvested for RNA isolation. Double-stranded cDNA was cloned unidirectionally into pSL1180. Inserts excised with EcoRI (5' end) and XhoI (3' end)."

ORIGIN

Alignment Scores:  
Pred. No.: 0.000251 Length: 806  
Score: 93.00 Matches: 17  
Percent Similarity: 80.00% Conservative: 3  
Best Local Similarity: 68.00% Mismatches: 5  
Query Match: 63.27% Indels: 0  
DB: 7 Gaps: 0

US-10-014-101B-32 (1-28) x CF393540 (1-806)

QY 4 GlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSerAlaMet 23  
|||||  
DB 335 GGACCATCTGCTATATCCGATCCTAGAACAAATGGGACTCTCGCATTCGCAATT 394  
|||||

QY 24 IleProGluIleAep 28  
|||||

DB 395 ATTCAGAGGAGGAC 409  
|||||

RESULT 5  
BM138354  
LOCUS  
DEFINITION  
library Triticum aestivum cDNA clone WHE0491\_F05\_K09, mRNA sequence.  
BM138354  
VERSION  
BM138354.1 GI:17147121  
KEYWORDS  
EST.  
SOURCE  
Triticum aestivum (bread wheat)  
ORGANISM  
Triticum aestivum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.

REFERENCE 1 (bases 1 to 539)  
 AUTHORS Anderson,O.D., Chao,S., Han,P.S., Heinen,S., Hsia,C.C., Kang,Y., Kruger,W.M., Lazo,G.R., Miller,S., Muehlbauer,G.J., Miller,R., Pritsch,C., Rausch,C.J., Seaton,C.L., Tong,J.C., Vance,C. and Wilson,C.F.  
 TITLE The structure and function of the expressed portion of the wheat genomes - Fusarium graminearum infected spike cDNA library  
 JOURNAL Unpublished (2001)  
 COMMENT US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center  
 800 Buchanan Street, Albany, CA 94710, USA  
 Tel: 5105595773  
 Fax: 5105595818  
 Email: canders@pw.usda.gov  
 Sequences have been trimmed to remove vector sequence and low quality sequence with phred score less than 20. No effort was taken to identify ESTs of fungal origin from this library, thus this EST could be of wheat or fungal origin.  
 Seq primer: Stratagene SK primer.

FEATURES  
 source  
 1..539  
 /organism="Triticum aestivum"  
 /mol\_type="mRNA"  
 /cultivar="Sumai3"  
 /db\_xref="taxon:4565"  
 /clone="WHE0491.F05.K09"  
 /tissue\_type="Spikes"  
 /dev\_stage="Adult plant"  
 /lab\_host="E. coli SOUR"  
 /clone\_lib="Wheat Fusarium graminearum infected spike cDNA library"  
 /notes="Vector: Lambda Uni-ZAP XR, excised phagemid pBluescript SK; Site\_1: EcoRI; Site\_2: XhoI; Plants were grown in the greenhouse. Spikes were sprayed at anthesis with Fusarium graminearum. Total RNA, and poly(A) RNA were prepared and pooled from infected spike at 0, 6, 12, 24, 36 and 48 hours after inoculation, a cDNA library was made, and the cDNA clones were in vivo excised to give pBluescript phagemids in G. Muehlbauer lab at the University of Minnesota (Kruger, W.M., Muehlbauer, G.J., Pritsch, C., Vance, C.). The cDNA library should contain genes of both wheat and fungal pathogen origin. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 0.000222 Length: 539  
 Score: 92.00 Matches: 17  
 Percent Similarity: 76.92% Conservative: 3  
 Best Local Similarity: 65.38% Mismatches: 6  
 Query Match: 62.59% Indels: 0  
 DB: 4 Gaps: 0

US-10-014-101B-32 (1-28) x BM138354 (1-539)

QY 1 SerAlaSerGlyLeuAlaLeuLeuTyrProThrAsnArgAsnLysTrpAspAsnArgMet 20  
 ||| :|||  
 Db 237 AGCAACAAATGGCCCATCTCTCTACCCAGTGAACAGATCAAGTGGGACACAGGACG 296  
 ||| :|||  
 QY 21 SerAlaMetIleProGlu 26  
 ||| :|||  
 Db 297 TCAGTGTGTATACCGGAG 314  
 ||| :|||

RESULT 6  
 CF398227  
 LOCUS  
 DEFINITION RTDS3\_24\_F01.b1\_A022 Drought-stressed loblolly pine roots DS3 Pinus taeda cDNA clone RTDS3\_24\_F01\_A022 3', mRNA sequence.  
 ACCESSION CF398227  
 VERSION CF398227.1 GI:34356644  
 KEYWORDS EST.

SOURCE Pinus taeda (loblolly pine)  
 ORGANISM Pinus taeda  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferales; Pinaceae; Pinus; Pinus.  
 REFERENCE 1 (bases 1 to 697)  
 AUTHORS Pratt,L., Cordonnier-Pratt,M.-M., Lorenz,W.W., Dean,J., Gebremedhin,M., Dervinis,C., Martin,T., White,T., Davis,J. and Neale,D.  
 TITLE An EST database from drought-stressed loblolly pine (Pinus taeda) roots  
 JOURNAL Unpublished (2003)  
 COMMENT Other ESTs: RTDS3\_24\_F01.g1\_A022  
 Contact: Cordonnier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmpratt@uga.edu  
 RNA prepared and library constructed by W. Walter Lorenz, School of Forestry, University of Georgia; plant material prepared at the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.  
 Seq primer: M13-21 (TGTAACGACGCGCAGT)  
 POLYA=Yes.

FEATURES  
 Location/Qualifiers  
 1..697  
 /organism="Pinus taeda"  
 /mol\_type="mRNA"  
 /strain="CCIONES"  
 /db\_xref="taxon:3352"  
 /clone="RTDS3\_24\_F01\_A022"  
 /lab\_host="DH10B-T1 phage-resistant E. coli"  
 /clone\_lib="Drought-stressed loblolly pine roots DS3"  
 /note="Vector: pSL1180; Site 1: EcoRI; Site 2: XhoI; The library was prepared from polyA+ RNA from drought-stressed loblolly pine (Pinus taeda) roots. Water was withheld from ramet clones until predawn needle water potential reached -1.75 MPa. On day 7 roots were harvested for RNA isolation. Double-stranded cDNA was cloned unidirectionally into pSL1180. Inserts excised with EcoRI (5' end) and XhoI (3' end)."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 0.000442 Length: 697  
 Score: 91.00 Matches: 16  
 Percent Similarity: 76.00% Conservative: 3  
 Best Local Similarity: 64.00% Mismatches: 6  
 Query Match: 61.90% Indels: 0  
 DB: 7 Gaps: 0

US-10-014-101B-32 (1-28) x CF398227 (1-697)

QY 4 GlyLeuAlaLeuLeuTyrProThrAsnArgAsnLysTrpAspAsnArgMetSerAlaMet 23  
 ||| :|||  
 Db 203 GGGCCGGTCTTATCTACCTGTGACCCCAATAGTGGGATAGCAATGTCAGCGGTG 262  
 ||| :|||  
 QY 24 IleProGluIleAsp 28  
 ||| :|||  
 Db 263 ACTCCTTAATGAGGAC 277  
 ||| :|||

RESULT 7  
 AV810994/c  
 LOCUS  
 DEFINITION AV810994 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-66-D02 3', mRNA sequence.  
 ACCESSION AV810994  
 VERSION AV810994.1 GI:19844979  
 KEYWORDS EST.

**SOURCE** Arabidopsis thaliana (thale cress)  
**ORGANISM** Arabidopsis thaliana  
**REFERENCE** Arabidopsis thaliana  
**AUTHORS** Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 401)  
 Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M., Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y. and Shinozaki, K.  
**TITLE** Large scale analysis of Arabidopsis full-length cDNA (2002b)  
**JOURNAL** Unpublished (2002)  
**COMMENT** Contact: Motoaki Seki  
 Plant Functional Genomics Research Group  
 RIKEN Genomic Sciences Center  
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
 Tel: 81-298-36-4359  
 Fax: 81-298-36-9060  
 Email: mseki@rtc.riken.go.jp  
 An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pluscript vector. Please visit our web site ([http://www.gsc.riken.go.jp/e/plant/index\\_e.html](http://www.gsc.riken.go.jp/e/plant/index_e.html)) for further details.

**FEATURES** Location/Qualifiers  
 source  
 1..401  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /clone="RAFL09-66-D02"  
 /dev\_stage="plants at various developmental stages from germination to mature seeds"  
 /lab\_host="DH10B"  
 /clone\_lib="RAFL9"  
 /notes="Site 1: BamHI; Site 2: SalI; subjected to dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24 hr) treatments"

**ORIGIN**  
 Alignment Scores:  
 Pred. No.: 0.000685 Length: 401  
 Score: 88.00 Matches: 15  
 Percent Similarity: 75.00% Conservative: 6  
 Best Local Similarity: 53.57% Mismatches: 7  
 Query Match: 59.86% Indels: 0  
 DB: 1 Gaps: 0

US-10-014-101B-32 (1-28) x AV810994 (1-401)  
 QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20  
 Db 360 TCACACTCTGCTGTTACTCTCTTTATCCACCACCGAACCAATGGACACCCCAATT 301  
 QY 21 SerAlaMetIleProGluIleAsp 28  
 Db 300 TCAACGATGACACCCGACGAGAT 277

**RESULT 8**  
 BI927011  
 LOCUS  
 DEFINITION EST546900 tomato flower, buds 0-3 mm Lycopersicon esculentum cDNA  
 clone cTOA31010 5' end, mRNA sequence.  
 ACCESSION BI927011  
 VERSION BI927011.1 GI:16236104  
 KEYWORDS EST.  
 SOURCE Lycopersicon esculentum (tomato)  
 ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon. 1 (bases 1 to 505)

**REFERENCE** 1 (bases 1 to 401)  
 Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M., Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y. and Shinozaki, K.  
 Large scale analysis of Arabidopsis full-length cDNA (2002b)  
 Unpublished (2002)  
 Contact: Motoaki Seki  
 Plant Functional Genomics Research Group  
 RIKEN Genomic Sciences Center  
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
 Tel: 81-298-36-4359  
 Fax: 81-298-36-9060  
 Email: mseki@rtc.riken.go.jp  
 An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pluscript vector. Please visit our web site ([http://www.gsc.riken.go.jp/e/plant/index\\_e.html](http://www.gsc.riken.go.jp/e/plant/index_e.html)) for further details.

**SOURCE** van der Hoeven, R.S., Bezzerides, J.L., Karamycheva, S.A., Tsai, J., Utterback, T., Van Aken, S., Ronning, C.M., Nierman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.  
**ORGANISM** Generation of ESTs from tomato flower tissue, 0-3 mm buds (2001)  
 Unpublished (2001)  
 Contact: CUGI  
 Clemson University Genomics Institute  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 This clone is available through the Clemson University Genomics Institute  
 Seq primer: T3  
 Location/Qualifiers  
 source  
 1..505  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="cTOA31010"  
 /tissue\_type="flower"  
 /dev\_stage="0-3mm buds"  
 /clone\_lib="tomato flower, buds 0-3 mm"  
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Cornell University; Flower buds and flowers were taken from greenhouse plants (4-8 wks old, TA496). They were immediately frozen in liquid nitrogen and then size-separated while remaining frozen."

**ORIGIN**  
 Alignment Scores:  
 Pred. No.: 0.000909 Length: 505  
 Score: 88.00 Matches: 15  
 Percent Similarity: 87.50% Conservative: 6  
 Best Local Similarity: 62.50% Mismatches: 3  
 Query Match: 59.86% Indels: 0  
 DB: 4 Gaps: 0

US-10-014-101B-32 (1-28) x BI927011 (1-505)  
 QY 3 SerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22  
 Db 81 AATGGACCATCTTGTCTACCCACACAGTAGGAAGATGGATAGGATGTAGCA 140  
 QY 23 MetIleProGlu 26  
 Db 141 ATAATACCAAGAA 152

**RESULT 9**  
 BI926735/c  
 LOCUS  
 DEFINITION EST546624 tomato flower, buds 0-3 mm Lycopersicon esculentum cDNA  
 clone cTOA30F5 5' end, mRNA sequence.  
 ACCESSION BI926735  
 VERSION BI926735.1 GI:16235507  
 KEYWORDS EST.  
 SOURCE Lycopersicon esculentum (tomato)  
 ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon. 1 (bases 1 to 642)  
 van der Hoeven, R.S., Bezzerides, J.L., Karamycheva, S.A., Tsai, J., Utterback, T., Van Aken, S., Ronning, C.M., Nierman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.  
 Generation of ESTs from tomato flower tissue, 0-3 mm buds (2001)  
 Unpublished (2001)  
 Contact: CUGI  
 Clemson University Genomics Institute  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>

This clone is available through the Clemson University Genomics Institute

Seq primer: T3.

FEATURES  
source  
1. .642  
Location/Qualifiers  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultiivar="TA496"  
/db\_xref="taxon:4081"  
/clone="CTOA30FS"  
/tissue\_type="flower"  
/dev\_stage="0-3mm buds"  
/clone\_lib="tomato flower, buds 0-3 mm"  
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Cornell University; sequencing: The Institute for Genomic Research; flower buds and flowers were taken from greenhouse plants (4-8 wks old, TA496). They were immediately frozen in liquid nitrogen and then size-separated while remaining frozen."

Alignment Scores:  
Pred. No.: 0.00122 Length: 642  
Score: 88.00 Matches: 15  
Percent Similarity: 87.50% Conservative: 6  
Best Local Similarity: 62.50% Mismatches: 3  
Query Match: 59.86% Indels: 0  
DB: 4 Gaps: 0

#### ORIGIN

US-10-014-101B-32 (1-28) x B1926735 (1-642)  
QY 3 SerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22  
Db 630 AATGGACCCACTTGTCTACCAAGTAGGAAAAGATGGATGATAGGATGTCAGCA 571  
QY 23 MetIleProGlu 26  
Db 570 ATAATACCAGAA 559

RESULT 10  
COL19329/c  
LOCUS  
DEFINITION GR\_Eb022C22.f GR\_Eb Gossypium raimondii cDNA clone GR\_Eb022C22  
5', mRNA sequence.

ACCESSION  
COL19329  
VERSION  
COL19329.1 GI:48818016

KEYWORDS  
EST.  
SOURCE  
Gossypium raimondii

ORGANISM  
Gossypium raimondii  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.

REFERENCE  
1 (bases 1 to 757)  
Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,  
Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and  
Wing,R.A.

TITLE  
Global assembly of Cotton ESTs

JOURNAL  
COMMENT  
Unpublished (2004)  
Contact: Rod A. Wing  
Arizona Genomics Institute  
The University of Arizona  
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA  
Tel: 520 626 9595  
Fax: 520 621 1259  
Email: http://genome.arizona.edu  
Plate: 022 row: C column: 22.

FEATURES  
source

Location/Qualifiers  
1. .757  
/organism="Gossypium raimondii"  
/mol\_type="mRNA"  
/db\_xref="taxon:29730"  
/clone="GR\_Eb022C22"  
/tissue\_type="floral"

/dev\_stage="3 to +3 DPA"  
/lab\_host="DH10B"  
/clone\_lib="GR\_Eb"  
/note="Vector: pCMV SPORT-6.1; Site 1: NotI; Site 2:  
EcoRV; Library made by Invitrogen with RNA supplied by  
Wendle lab. Directional cloned into NotI-EV. Colonies  
plated/picked by AGI. More glycerol clones held in -80."

#### ORIGIN

Alignment Scores:  
Pred. No.: 0.00149 Length: 757  
Score: 88.00 Matches: 15  
Percent Similarity: 83.33% Conservative: 5  
Best Local Similarity: 62.50% Mismatches: 4  
Query Match: 59.86% Indels: 0  
DB: 7 Gaps: 0

US-10-014-101B-32 (1-28) x COL19329 (1-757)

QY 3 SerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22  
Db 545 ACAGGCCCTGCTCTCGTTTATCTATGAACACACAAAAGTGGATGATAGGATGTCAGCT 486

QY 23 MetIleProGlu 26

Db 485 GTAATACCAGAT 474

#### RESULT 11

COL1913/c

LOCUS  
DEFINITION GR\_Eb01P23.r GR\_Eb Gossypium raimondii cDNA clone GR\_Eb01P23',  
mRNA sequence.

ACCESSION  
COL1913

VERSION  
COL1913.1 GI:48816600

KEYWORDS  
EST.

SOURCE  
Gossypium raimondii

ORGANISM  
Gossypium raimondii  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.

REFERENCE  
1 (bases 1 to 795)

Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,  
Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and  
Wing,R.A.

TITLE  
Global assembly of Cotton ESTs

JOURNAL  
COMMENT  
Unpublished (2004)  
Contact: Rod A. Wing  
Arizona Genomics Institute  
The University of Arizona  
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA  
Tel: 520 626 9595  
Fax: 520 621 1259  
Email: http://genome.arizona.edu  
Plate: 01 row: P column: 23.

FEATURES  
source

Location/Qualifiers  
1. .795  
/organism="Gossypium raimondii"  
/mol\_type="mRNA"  
/db\_xref="taxon:29730"  
/clone="GR\_Eb01P23"  
/tissue\_type="floral"  
/dev\_stage="3 to +3 DPA"  
/lab\_host="DH10B"  
/clone\_lib="GR\_Eb"

/note="Vector: pCMV SPORT-6.1; Site 1: NotI; Site 2:  
EcoRV; Library made by Invitrogen with RNA supplied by  
Wendle lab. Directional cloned into NotI-EV. Colonies  
plated/picked by AGI. More glycerol clones held in -80."

#### ORIGIN

Alignment Scores:  
Pred. No.: 0.00159 Length: 795  
Score: 88.00 Matches: 15

Percent Similarity: 83.33%  
 Best Local Similarity: 62.50%  
 Query Match: 59.86%  
 DB: 7

US-10-014-101B-32 (1-28) x CO117913 (1-795)  
 QY 3 SerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22  
 Db 518 ACAGGACCTGCTCGTTATCTATGACAGACAAAGTGGATGATAGATGTCAGCT 459  
 QY 23 MetIleProGlu 26  
 Db 458 GTAATACCAGAT 447

## RESULT 12

CA766652

## LOCUS

DEFINITION CA766652 807 bp mRNA linear EST 08-JAN-2003  
 Oryza sativa (indica cultivar-group) cDNA clone C0004334 5' similar to Hypothetical 47.9 kDa oxidoreductase in fasciation locus (ORF5), mRNA sequence.

## ACCESSION

CA766652

## VERSION

CA766652.2

## KEYWORDS

GI:27548636

## SOURCE

EST.

## ORGANISM

Oryza sativa (indica cultivar-group)

Oryza sativa (indica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzeae; Oryza.

1 (bases 1 to 807)

Bennett, J., Arumugam, K., Lafitte, R., Wen, J., Rudd, S. and  
 Bruskiewicz, R.M.

IRRI Drought Stress Panicle cDNA Library

Unpublished (2002)

On Dec 2, 2002 this sequence version replaced gi:25995907.

Contact: Richard Bruskiewicz

Biometrics and Bioinformatics Unit

International Rice Research Institute

DAPO 7777, Metro Manila, Philippines

Tel: +63-2-845-0563

Fax: +63-2-845-0606

Email: r.bruskiewicz@cgiar.org

International Rice Information System (IRIS);  
 http://www.iris.irri.org: D0204333Assignment of putative function to the sequence by S. Rudd of the  
 Munich Information Center for Protein Sequences  
 (http://mips.gsf.de)

Plate: 12 row: E column: 14.

## FEATURES

source

1..807

Location/Qualifiers

/organism="Oryza sativa (indica cultivar-group)"

/mol\_type="mRNA"

/cultivar="IR64"

/db\_xref="taxon:39946"

/clone="C0004334"

/tissue\_type="Panicles"

/dev\_stage="Flowering"

/clone\_lib="IRRI Drought Stress Panicle Library"

/notes="Vector: pBluescript II SK+; Water stress was  
 applied by not watering for 4 consecutive days. Panicles  
 were collected from control (well watered) and stressed  
 plants at 2 days before heading, at heading, 50% flowering  
 and 4 days after 50% flowering."

## ORIGIN

Alignment Scores:

Pred. No.: 0.00161

Score: 88.00

Percent Similarity: 73.08%

Best Local Similarity: 61.54%

Query Match: 59.86%

DB: 6

Length: 807

Matches: 16

Conservative: 3

Mismatches: 7

Indels: 0

Gaps: 0

Conservative: 5  
 Mismatches: 4  
 Indels: 0  
 Gaps: 0

US-10-014-101B-32 (1-28) x CA766652 (1-807)

QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMet 20  
 Db 532 AGCAACAATGGTCCCATACTGCTTTACCCAGTCAGTACAGACCAAGTGGGACACAGACA 591

QY 21 SerAlaMetIleProGlu 26

Db 592 TCAGTGGTCATACCAGAT 609

## RESULT 13

CO121610

## LOCUS

DEFINITION

CO121610

ACCESSION

CO121610

VERSION

CO121610.1

KEYWORDS

EST.

SOURCE

Gossypium raimondii

Gossypium raimondii

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.

1 (bases 1 to 824)

Kim, H., Yu, Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C.,  
 Udall, J.A., Rapp, R.A., Wendel, J.F., Rao, K., Soderlund, C. and  
 Wing, R.A.

Global assembly of Cotton ESTs

Unpublished (2004)

Contact: Rod A. Wing

Arizona Genomics Institute

The University of Arizona

Forbes Building Room 303, Tucson, AZ, 85721-0036, USA

Tel: 520 626 9595

Fax: 520 621 1259

Email: http://genome.arizona.edu

Plate: 02 row: N column: 24.

## FEATURES

source

1..824

Location/Qualifiers

/organism="Gossypium raimondii"

/mol\_type="mRNA"

/db\_xref="taxon:29730"

/clone="GR\_Eb02N24"

/tissue\_type="floral"

/dev\_stage="3 to +3 DPA"

/lab\_host="DH10B"

/clone\_lib="GR\_Eb"

/note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:  
 EcoRV; Library made by invitrogen with RNA supplied by  
 Wendle lab. Directional cloned into NotI-EV. Colonies  
 plated/picked by AGI. More glycerol clones held in -80."

## ORIGIN

Alignment Scores:

Pred. No.: 0.00166

Score: 88.00

Percent Similarity: 83.33%

Best Local Similarity: 62.50%

Query Match: 59.86%

DB: 7

Length: 824

Matches: 15

Conservative: 5

Mismatches: 4

Indels: 0

Gaps: 0

US-10-014-101B-32 (1-28) x CO121610 (1-824)

QY 3 SerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22

Db 735 ACAGGACCTGCTCGTTATCTATGACAGACAAAGTGGATGATAGATGTCAGCT 794

QY 23 MetIleProGlu 26

Db 795 GTAATACCAGAT 806

## RESULT 14

Unpublished (2003)  
Contact: Rod Wing  
Arizona Genomics Institute  
University of Arizona  
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ  
85721-0088, USA  
Tel: 520 626 3967  
Fax: 520 621 9288  
Email: http://genome.arizona.edu

PCR Primers  
FORWARD: gta aaa cga cgg cca gtc  
BACKWARD: gga aac agc tat gac cat g  
Plate: 05 row: N column: 03  
Seq primer: gta aaa cga cgg cca gtc.

location/Qualifiers  
1..772  
/organism="Oryza sativa (indica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="IR36"  
/db\_xref="taxon:39946"  
/clone="OSIIEa05N03"  
/tissue\_type="leaf"  
/dev\_stage="3 week"  
/lab\_host="DH10B"  
/clone\_lib="OSIIEa"  
/note="Vector: pBluescript II KS +; Site\_1: EcoRI; Site\_2: XhoI; Lesion Mimic SPL 11"

ORIGIN  
Alignment Scores:  
Pred. No.: 0.00222 Length: 772  
Score: 87.00 Matches: 16  
Percent Similarity: 74.07% Conservative: 4  
Best Local Similarity: 59.26% Mismatches: 7  
Query Match: 59.18% Indels: 0  
DB: 6 Gaps: 0

US-10-014-101B-32 (1-28) x CB620527 (1-772)  
QY 2 AlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSer 21  
DB 498 GCCATGGGCCCCGTCCTCATCTACCCATGACCGACACGACAGTGGGACAGTACATGTCG 557  
QY 22 AlaMetIleProGluIleAsp 28  
DB 558 GCAGTGATCAGCGACGACGAC 578

RESULT 15  
LOCUS CO435706 863 bp mRNA linear EST 01-OCT-2004  
DEFINITION OSMR772 5MT resistant rice mutant cDNA library Oryza sativa (japonica cultivar-group) cDNA, mRNA sequence.  
ACCESSION CO435706  
VERSION CO435706.1 GI:53553154  
KEYWORDS EST.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.  
1 (bases 1 to 863)  
Kim,D.S., Lee,I.S., Kang,S.Y., Song,H.S., Lee,Y.I. and Seo,Y.W.  
Expressed sequence tag (EST) analysis and characterization of differentially expressed genes related to 5-methyltryptophan (5MT) resistance in mutant rice  
Unpublished (2003)  
Contact: Seo YW  
Department of Crop Science, division of Biotechnology and Genetic Engineering  
Korea University  
Anam-Dong, Seongbuk-Gu, Seoul, 136-701, Korea  
Tel: +82 2 3290 3005  
Fax: +82 2 3290 3501

JOURNAL COMMENT  
Unpublished (2003)  
Contact: Rod Wing  
Arizona Genomics Institute  
University of Arizona  
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ  
85721-0088, USA  
Tel: 520 626 3967  
Fax: 520 621 9288  
Email: http://genome.arizona.edu

PCR Primers  
FORWARD: gta aaa cga cgg cca gtc  
BACKWARD: gga aac agc tat gac cat g  
Plate: 05 row: N column: 03  
Seq primer: gta aaa cga cgg cca gtc.

location/Qualifiers  
1..772  
/organism="Oryza sativa (indica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="IR36"  
/db\_xref="taxon:39946"  
/clone="OSIIEa05N03"  
/tissue\_type="leaf"  
/dev\_stage="3 week"  
/lab\_host="DH10B"  
/clone\_lib="OSIIEa"  
/note="Vector: pBluescript II KS +; Site\_1: EcoRI; Site\_2: XhoI; Lesion Mimic SPL 11"

ORIGIN  
Alignment Scores:  
Pred. No.: 0.00222 Length: 772  
Score: 87.00 Matches: 16  
Percent Similarity: 74.07% Conservative: 4  
Best Local Similarity: 59.26% Mismatches: 7  
Query Match: 59.18% Indels: 0  
DB: 6 Gaps: 0

US-10-014-101B-32 (1-28) x CB620527 (1-772)  
QY 2 AlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSer 21  
DB 498 GCCATGGGCCCCGTCCTCATCTACCCATGACCGACACGACAGTGGGACAGTACATGTCG 557  
QY 22 AlaMetIleProGluIleAsp 28  
DB 558 GCAGTGATCAGCGACGACGAC 578

RESULT 15  
LOCUS CO435706 863 bp mRNA linear EST 01-OCT-2004  
DEFINITION OSMR772 5MT resistant rice mutant cDNA library Oryza sativa (japonica cultivar-group) cDNA, mRNA sequence.  
ACCESSION CO435706  
VERSION CO435706.1 GI:53553154  
KEYWORDS EST.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.  
1 (bases 1 to 863)  
Kim,D.S., Lee,I.S., Kang,S.Y., Song,H.S., Lee,Y.I. and Seo,Y.W.  
Expressed sequence tag (EST) analysis and characterization of differentially expressed genes related to 5-methyltryptophan (5MT) resistance in mutant rice  
Unpublished (2003)  
Contact: Seo YW  
Department of Crop Science, division of Biotechnology and Genetic Engineering  
Korea University  
Anam-Dong, Seongbuk-Gu, Seoul, 136-701, Korea  
Tel: +82 2 3290 3005  
Fax: +82 2 3290 3501

JOURNAL COMMENT  
Unpublished (2003)  
Contact: Rod Wing  
Arizona Genomics Institute  
University of Arizona  
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ  
85721-0088, USA  
Tel: 520 626 3967  
Fax: 520 621 9288  
Email: http://genome.arizona.edu

PCR Primers  
FORWARD: gta aaa cga cgg cca gtc  
BACKWARD: gga aac agc tat gac cat g  
Plate: 05 row: N column: 03  
Seq primer: gta aaa cga cgg cca gtc.

location/Qualifiers  
1..772  
/organism="Oryza sativa (indica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="IR36"  
/db\_xref="taxon:39946"  
/clone="OSIIEa05N03"  
/tissue\_type="leaf"  
/dev\_stage="3 week"  
/lab\_host="DH10B"  
/clone\_lib="OSIIEa"  
/note="Vector: pBluescript II KS +; Site\_1: EcoRI; Site\_2: XhoI; Lesion Mimic SPL 11"

ORIGIN  
Alignment Scores:  
Pred. No.: 0.00222 Length: 772  
Score: 87.00 Matches: 16  
Percent Similarity: 74.07% Conservative: 4  
Best Local Similarity: 59.26% Mismatches: 7  
Query Match: 59.18% Indels: 0  
DB: 6 Gaps: 0

US-10-014-101B-32 (1-28) x CB620527 (1-772)  
QY 2 AlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSer 21  
DB 498 GCCATGGGCCCCGTCCTCATCTACCCATGACCGACACGACAGTGGGACAGTACATGTCG 557  
QY 22 AlaMetIleProGluIleAsp 28  
DB 558 GCAGTGATCAGCGACGACGAC 578

RESULT 15  
LOCUS CO435706 863 bp mRNA linear EST 01-OCT-2004  
DEFINITION OSMR772 5MT resistant rice mutant cDNA library Oryza sativa (japonica cultivar-group) cDNA, mRNA sequence.  
ACCESSION CO435706  
VERSION CO435706.1 GI:53553154  
KEYWORDS EST.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.  
1 (bases 1 to 863)  
Kim,D.S., Lee,I.S., Kang,S.Y., Song,H.S., Lee,Y.I. and Seo,Y.W.  
Expressed sequence tag (EST) analysis and characterization of differentially expressed genes related to 5-methyltryptophan (5MT) resistance in mutant rice  
Unpublished (2003)  
Contact: Seo YW  
Department of Crop Science, division of Biotechnology and Genetic Engineering  
Korea University  
Anam-Dong, Seongbuk-Gu, Seoul, 136-701, Korea  
Tel: +82 2 3290 3005  
Fax: +82 2 3290 3501

JOURNAL COMMENT  
Unpublished (2003)  
Contact: Rod Wing  
Arizona Genomics Institute  
University of Arizona  
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ  
85721-0088, USA  
Tel: 520 626 3967  
Fax: 520 621 9288  
Email: http://genome.arizona.edu

PCR Primers  
FORWARD: gta aaa cga cgg cca gtc  
BACKWARD: gga aac agc tat gac cat g  
Plate: 05 row: N column: 03  
Seq primer: gta aaa cga cgg cca gtc.

location/Qualifiers  
1..772  
/organism="Oryza sativa (indica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="IR36"  
/db\_xref="taxon:39946"  
/clone="OSIIEa05N03"  
/tissue\_type="leaf"  
/dev\_stage="3 week"  
/lab\_host="DH10B"  
/clone\_lib="OSIIEa"  
/note="Vector: pBluescript II KS +; Site\_1: EcoRI; Site\_2: XhoI; Lesion Mimic SPL 11"

ORIGIN  
Alignment Scores:  
Pred. No.: 0.00222 Length: 772  
Score: 87.00 Matches: 16  
Percent Similarity: 74.07% Conservative: 4  
Best Local Similarity: 59.26% Mismatches: 7  
Query Match: 59.18% Indels: 0  
DB: 6 Gaps: 0

US-10-014-101B-32 (1-28) x CB620527 (1-772)  
QY 2 AlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSer 21  
DB 498 GCCATGGGCCCCGTCCTCATCTACCCATGACCGACACGACAGTGGGACAGTACATGTCG 557  
QY 22 AlaMetIleProGluIleAsp 28  
DB 558 GCAGTGATCAGCGACGACGAC 578

RESULT 15  
LOCUS CO435706 863 bp mRNA linear EST 01-OCT-2004  
DEFINITION OSMR772 5MT resistant rice mutant cDNA library Oryza sativa (japonica cultivar-group) cDNA, mRNA sequence.  
ACCESSION CO435706  
VERSION CO435706.1 GI:53553154  
KEYWORDS EST.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.  
1 (bases 1 to 863)  
Kim,D.S., Lee,I.S., Kang,S.Y., Song,H.S., Lee,Y.I. and Seo,Y.W.  
Expressed sequence tag (EST) analysis and characterization of differentially expressed genes related to 5-methyltryptophan (5MT) resistance in mutant rice  
Unpublished (2003)  
Contact: Seo YW  
Department of Crop Science, division of Biotechnology and Genetic Engineering  
Korea University  
Anam-Dong, Seongbuk-Gu



RESULT 20	BI203744	642 bp	mRNA	linear	EST 11-JUL-2001
LOCUS	EST521784	ctos	Lycopersicon	esculentum	cdna clone cTOS2N10 5' end,
DEFINITION	mRNA sequence.				
ACCESSION	BI203744				
KEYWORDS	BI203744.1	GI:14681468			
SOURCE	EST.				
ORGANISM	Lycopersicon	esculentum	(tomato)		
REFERENCE	Lycopersicon esculentum				
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.				
TITLE	1 (bases 1 to 642)				
JOURNAL	van der Hoeven, R., Sun, H., Bezzerides, J., Cho, J., Utterback, R., Ronning, C. and Tanksley, S.				
COMMENT	Generation of ESTs from Tomato Suspension Cultures Unpublished (2001)				
FEATURES	Contact: CUGI				
source	Clemson University Genomics Institute				
	Clemson University				
	100 Jordan Hall, Clemson, SC 29634, USA				
	Email: <a href="http://www.genome.clemson.edu/orders/index.html">http://www.genome.clemson.edu/orders/index.html</a> .				
	Location/Qualifiers				
	1. .642				
	/organism="Lycopersicon esculentum"				
	/mol_type="mRNA"				
	/cultivar="TA96, E6203"				
	/db_xref="taxon:4081"				
	/clone="ctos2N10"				
	/tissue_type="suspension cultures"				
	/lab_host="SOLR"				
	/clone_lib="ctos"				
	/note=Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; Suspension cultures of L.esculentum E6203 were grown in Murashige and Skoog based medium, supplemented with 15% coconut milk (filter sterilized and added after autoclaving), 2% saccharose, and 1mg/ml 2,4D (pH5.8). Fresh medium was added every 7 days, and cultures were grown at 25 C, with 12hrs of light and continuous shaking."				
ORIGIN					
Alignment Scores:					
Pred. No.:	0.00257	Length:	642		
Score:	86.00	Matches:	15		
Percent Similarity:	83.33%	Conservative:	5		
Best Local Similarity:	62.50%	Mismatches:	4		
Query Match:	58.50%	Indels:	0		
DB:	4	Gaps:	0		
US-10-014-101B-32	(1-28) x BI203744	(1-642)			
QY	3	SerGlyLeuAlaLeuLeuTyrProThrAsnArgAsnLysTrpAspAsnArgMetSerAla	22		
Db	114	ACTGGACCCATCTGTTCTTACCCACAAAGTCGGAAGATGGATGATGATGTCAGCC	173		
QY	23	MetIleProGlu	26		
Db	174	ACGATACCCAGAA	185		
RESULT 21					
LOCUS	BI204708	715 bp	mRNA	linear	EST 11-JUL-2001
DEFINITION	EST522748				
ACCESSION	BI204708	ctos	Lycopersicon	esculentum	cdna clone cTOS5H12 5' end,
KEYWORDS	BI204708.1	GI:14682432			
SOURCE	EST.				
ORGANISM	Lycopersicon	esculentum	(tomato)		
REFERENCE	Lycopersicon esculentum				
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;				
TITLE	1 (bases 1 to 715)				
JOURNAL	van der Hoeven, R., Sun, H., Bezzerides, J., Cho, J., Utterback, R., Ronning, C. and Tanksley, S.				
COMMENT	Generation of ESTs from Tomato Suspension Cultures Unpublished (2001)				
FEATURES	Contact: CUGI				
source	Clemson University Genomics Institute				
	Clemson University				
	100 Jordan Hall, Clemson, SC 29634, USA				
	Email: <a href="http://www.genome.clemson.edu/orders/index.html">http://www.genome.clemson.edu/orders/index.html</a> .				
	Location/Qualifiers				
	1. .715				
	/organism="Lycopersicon esculentum"				
	/mol_type="mRNA"				
	/cultivar="TA96, E6203"				
	/db_xref="taxon:4081"				
	/clone="ctos2N10"				
	/tissue_type="suspension cultures"				
	/lab_host="SOLR"				
	/clone_lib="ctos"				
	/note=Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; Suspension cultures of L.esculentum E6203 were grown in Murashige and Skoog based medium, supplemented with 15% coconut milk (filter sterilized and added after autoclaving), 2% saccharose, and 1mg/ml 2,4D (pH5.8). Fresh medium was added every 7 days, and cultures were grown at 25 C, with 12hrs of light and continuous shaking."				
ORIGIN					
Alignment Scores:					
Pred. No.:	0.00161	Length:	438		
Score:	86.00	Matches:	15		
Percent Similarity:	76.92%	Conservative:	5		
Best Local Similarity:	57.69%	Mismatches:	6		
Query Match:	58.50%	Indels:	0		
DB:	4	Gaps:	0		
US-10-014-101B-32	(1-28) x BG518327	(1-438)			
QY	1	SerAlaSerGlyLeuAlaLeuLeuTyrProThrAsnArgAsnLysTrpAspAsnArgMet	20		
Db	228	AGCAACAATGGTCCCATATGCTTTATCCAGTGAACAAATCAAAGTCGGACACAGACG	287		
QY	21	SerAlaMetIleProGlu	26		
Db	288	TCAGTAGTCATACACAGAT	305		

```

asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 715)
REFERENCE
AUTHORS van der Hoeven,R., Sun,H., Bezzerides,J., Cho,J., Utterback,T.,
Ronning,C. and Tanksley,S.
TITLE Generation of ESTs from Tomato Suspension Cultures
JOURNAL Unpublished (2001)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.

FEATURES
source
1..715
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496, E6203"
/db_xref="taxon:4081"
/clone="CTOSH12"
/tissue_type="suspension cultures"
/lab_host="SOLR"
/clone_lib="CTOS"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; Suspension cultures of L.esculentum E6203 were grown
in Murashige and Skoog based medium, supplemented with 1%
coconut milk (filter sterilized and added after
autoclaving), 2% saccharose, and 1mg/ml 2,4D (pH5.8).
Fresh medium was added every 7 days, and cultures were
grown at 25 C, with 12hrs of light and continuous
shaking."

ORIGIN
Alignment Scores:
Pred. No.: 0.00293 Length: 715
Score: 86.00 Matches: 15
Percent Similarity: 83.3% Conservative: 5
Best Local Similarity: 62.50% Mismatches: 4
Query Match: 58.50% Indels: 0
DB: 4 Gaps: 0

US-10-014-101B-32 (1-28) x B1204708 (1-715)

QY 3 SerGlyLeuAlaLeuLeuTyrProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 114 ACTGACCCACTTGTCTACCAACAAGTCGGAAGATGGATGATGATGTCAGCC 173
QY 23 MetIleProGlu 26
Db 174 ACGATACCAGAA 185

RESULT 22
B2644495/c 826 bp DNA linear GSS 29-JAN-2003
LOCUS OGCCW96TM_ZM_0.7_1.5_KB_Zea_mays_genomic_clone_ZM5BMA0143P23,
DEFINITION genomic survey sequence.
ACCESSION B2644495
VERSION B2644495.1 GI:28106659
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other GSSs: OGCCW96TC
Contact: Cathy Whitelaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843

FEATURES
source
1..826
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM5BMA0143P23"
/clone_lib="ZM_0.7_1.5_KB"
/notes="Vector: pBCK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Alignment Scores:
Pred. No.: 0.0035 Length: 826
Score: 86.00 Matches: 13
Percent Similarity: 80.77% Conservative: 8
Best Local Similarity: 50.00% Mismatches: 5
Query Match: 58.50% Indels: 0
DB: 8 Gaps: 0

US-10-014-101B-32 (1-28) x B2644495 (1-826)

QY 3 SerGlyLeuAlaLeuLeuTyrProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 326 GCGGGATCATCTCATGTACCCATGAACAGGACAGTGGGACGCGGATGACAGCG 267
QY 23 MetIleProGluIleAsp 28
Db 266 ATGACCCAGCCAGCGAC 249

RESULT 23
CG289642/c 906 bp DNA linear GSS 25-AUG-2003
LOCUS OGWFN30TH_ZM_0.7_1.5_KB_Zea_mays_genomic_clone_ZM5BMA0566F12,
DEFINITION genomic survey sequence.
ACCESSION CG289642
VERSION CG289642.1 GI:34203856
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 906)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other GSSs: OGWFN30TV
Contact: Cathy Whitelaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source
1..906
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM5BMA0566F12"
/clone_lib="ZM_0.7_1.5_KB"
/notes="Vector: pBCK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

```

```

Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source
1..826
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM5BMA0143P23"
/clone_lib="ZM_0.7_1.5_KB"
/notes="Vector: pBCK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Alignment Scores:
Pred. No.: 0.0035 Length: 826
Score: 86.00 Matches: 13
Percent Similarity: 80.77% Conservative: 8
Best Local Similarity: 50.00% Mismatches: 5
Query Match: 58.50% Indels: 0
DB: 8 Gaps: 0

US-10-014-101B-32 (1-28) x B2644495 (1-826)

QY 3 SerGlyLeuAlaLeuLeuTyrProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 326 GCGGGATCATCTCATGTACCCATGAACAGGACAGTGGGACGCGGATGACAGCG 267
QY 23 MetIleProGluIleAsp 28
Db 266 ATGACCCAGCCAGCGAC 249

RESULT 23
CG289642/c 906 bp DNA linear GSS 25-AUG-2003
LOCUS OGWFN30TH_ZM_0.7_1.5_KB_Zea_mays_genomic_clone_ZM5BMA0566F12,
DEFINITION genomic survey sequence.
ACCESSION CG289642
VERSION CG289642.1 GI:34203856
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 906)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other GSSs: OGWFN30TV
Contact: Cathy Whitelaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source
1..906
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM5BMA0566F12"
/clone_lib="ZM_0.7_1.5_KB"
/notes="Vector: pBCK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

```

```

Alignment Scores:
Pred. No.: 0.00392 Length: 906
Score: 86.00 Matches: 13
Percent Similarity: 80.77% Conservatives: 8
Best Local Similarity: 50.00% Mismatches: 5
Query Match: 58.50% Indels: 0
DB: 9 Gaps: 0

US-10-014-101B-32 (1-28) x CG289642 (1-906)

QY 3 SerGlyLeuAlaLeuLeuTyPProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22
Db 672 GCCGGATCATCTCTATGTTACCCATGACAAAGACAGGTGGGACGACCGATGACAGCG 613

QY 23 MetIleProGluIleAasp 28
Db 612 ATGACCCCGACCGGAC 595

RESULT 24
CG351153/c 963 bp DNA linear GSS 26-AUG-2003
LOCUS CG351153
DEFINITION O22BM20TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBma0753D15,
genomic survey sequence.
ACCESSION CG351153
VERSION CG351153.1 GI:34268419
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 963)
REFERENCE Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
AUTHORS Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nummer,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
TITLE Unpublished (2002)
JOURNAL Other GSSs: CG2BM20TH
COMMENT Contact: Cathy Whitelaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: Tg
Class: sheared ends.
FEATURES
source
location/Qualifiers
1..963
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0753D15"
/clone_lib="ZM_0.7_1.5_KB"
/notes="Vector: pBCKS; Site_1: HincII; 0:7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Alignment Scores:
Pred. No.: 0.00422 Length: 963
Score: 86.00 Matches: 13
Percent Similarity: 80.77% Conservatives: 8
Best Local Similarity: 50.00% Mismatches: 5
Query Match: 58.50% Indels: 0
DB: 9 Gaps: 0

US-10-014-101B-32 (1-28) x CG351153 (1-963)

QY 3 SerGlyLeuAlaLeuLeuTyPProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22
Db 487 GCCGGATCATCTCTATGTTACCCATGACAAAGACAGGTGGGACGACCGATGACAGCG 428

```

```

QY 23 MetIleProGluIleAasp 28
Db 427 ATGACCCCGACCGGAC 410

RESULT 25
CG3513923
LOCUS CG3513923
DEFINITION VVD172F06 377903 An expressed sequence tag database for abiotic
stressed berries of Vitis vinifera var. Chardonnay Vitis vinifera
cDNA clone VVD172F06 5, mRNA sequence.
*
ACCESSION CB913923
VERSION CB913923.1 GI:30128584
KEYWORDS EST.
SOURCE Vitis vinifera
ORGANISM Vitis vinifera
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; Vitaceae; Vitis.
1 (bases 1 to 584)
REFERENCE Cushman,J.C.
AUTHORS An expressed sequence tag database for abiotic stressed berries of
TITLE Vitis vinifera var. Chardonnay
JOURNAL Unpublished (2002)
COMMENT Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu
PCR Primers
FORWARD: T3 20mer
BACKWARD: T7 21mer (backward)
Plate: 172 row: F column: 06
Seq primer: T3 20mer
High quality sequence stop: 584.
FEATURES
source
location/Qualifiers
1..584
/organism="Vitis vinifera"
/mol_type="mRNA"
/db_xref="taxon:29760"
/clone="VVD172F06"
/tissue_type="berries"
/dev_stage="mixed; 8, 9, 11, 13, 15, 16 weeks daf"
/clone_lib="An expressed sequence tag database for abiotic
stressed berries of Vitis vinifera var. Chardonnay"
/notes="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site_1:
EcoRI; Site_2: XhoI"
ORIGIN
Alignment Scores:
Pred. No.: 0.00331 Length: 584
Score: 85.00 Matches: 15
Percent Similarity: 73.08% Conservatives: 4
Best Local Similarity: 57.63% Mismatches: 7
Query Match: 57.82% Indels: 0
DB: 6 Gaps: 0

US-10-014-101B-32 (1-28) x CB913923 (1-584)

QY 3 SerGlyLeuAlaLeuLeuTyPProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22
Db 97 AGTGGCCCATCTCTATGTTACCCATGACAAACAAAGTGGGACGACCGACATCGGTG 156

QY 23 MetIleProGluIleAasp 28
Db 157 GTTACCGCAGAGGAGGAC 174

RESULT 26
CG210007/c
LOCUS CG210007
DEFINITION CAB20005_IVb_Rb_G06 Cabernet Sauvignon Flower bloom - CAB2 Vitis
vinifera cDNA clone CAB20005_IVb_Rb_G06 3', mRNA sequence.

```

```

ACCESSION   CF210007
VERSION     CF210007.1
KEYWORDS    GI:33404380
SOURCE      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
SOURCE      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
SOURCE      rosids; Vitaceae; Vitis.
ORGANISM    Vitis vinifera
REFERENCE   1 (bases 1 to 741)
AUTHORS     Goes da Silva, F., Iandolino, A., Lim, H., Baek, J., Leslie, A., Xu, J.,
            Jones, K. and Cook, D.
TITLE       Expressed sequence tags from Vitis vinifera 'Cabernet Sauvignon'
            berries at various developmental stages
JOURNAL     Unpublished (2003)
COMMENT     Contact: Douglas Cook, PhD
            CAES Genome Facility
            UC Davis, Plant Pathology
            One Shields Ave, Davis, CA 95616, USA
            Tel: 530 754 6561
            Fax: 530 754 6617
            Email: drcocook@ucdavis.edu
            Seq primer: GCCAACGATGCTGTAG.
FEATURES    Location/Qualifiers
             1..741
                /organism="Vitis vinifera"
                /mol_type="mRNA"
                /cultivar="Cabernet Sauvignon"
                /db_xref="taxon:29760"
                /clone="CAB20005 IVb Rb_G06"
                /sex="Hermaphrodite"
                /dev_stage="Bloom"
                /clone_lib="Cabernet Sauvignon Flower bloom - CAB2"
                /notes="Organ: Flower - Bloom; Vector: pDNR; Site 1: SfiI;
                Site 2: SfiI; CAB2 is a cDNA library of Vitis vinifera cv.
                'Cabernet Sauvignon' Clone 8 berries. Samples were
                collected at full bloom (80 to 100% flowers showing
                dehiscence of calypters or caps and anthers fully
                extended). Sampled vines were located at the University of
                California, Davis, Experimental Vineyard. cDNAs were made
                by oligo-dT priming and directionally cloned. 5' and 3'
                adaptors were used in cloning as follows:
                5'-AACGATGGTATCAACGCGAGTGGCATTCAGCGCGG-3' and
                5'-ATTCTAGAGCCGAGCGCGGACATG-DT(30)NN-3'. Library was
                constructed using the Clontech Creator SMART kit and
                size-selected to contain the 0.5-3 kb size fraction."
ORIGIN
Alignment Scores:
Pred. No.: 0.00444 Length: 741
Score: 85.00 Matches: 15
Percent Similarity: 73.08% Conservative: 4
Best Local Similarity: 57.69% Mismatches: 7
Query Match: 57.82% Indels: 0
DB: 7 Gaps: 0

US-10-014-101B-32 (1-28) x CF210007 (1-741)
QY 3 SerGlyLeuAlaLeuLeuTyPProThrAsnArgAsnLysTrpAspSerAla 22
DB 501 AGTGGCCCCATCTCTATTACCTATGACAAACAAACAGTGGGACCGACATCGGTG 442
QY 23 MetileProGluLeuLeuP 28
DB 441 GTTACGCCAGGAGGAC 424

RESULT 27
BI970438/c
LOCUS      BI970438
DEFINITION GM630010B12B02 Gm-r1083 Glycine max cDNA clone Gm-r1083-3531 3',
            mRNA sequence.
ACCESSION  BI970438
VERSION    BI970438.1
KEYWORDS   GI:16344843

```

Glycine max (soybean)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

1 (bases 1 to 773)

Vodkin, L., Keim, P., Shoemaker, R., Retzel, E., Khanna, A., Coryell, V., Erpelnding, J., Raph, C., Shoop, E., Pardinas, J., Liu, B. and Lewin, H. A Functional Genomics Program for Soybean (NSF 9872565)

Unpublished (1999)

Other ESTs: AW830743 corresponding to Gm-cl028-5914 (5')

Contact: Vodkin, L.O., PI, A Functional Genomics Program for Soybean (NSF 9872565)

Lewin, H. A., Director, Keck Center for Comparative and Functional Genomics

University of Illinois

Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA

Tel: (217) 244-6147

Fax: (217) 333-4582

Email: l-vodkin@uiuc.edu

This clone is available through: Incyte Genomics, 4633 World Parkway Circle St. Louis, Missouri 63134. Phone (800) 430-0030 or (314) 427-3222 FAX: (314) 427-3324. Web site: <http://www.incyte.com/reagents/catalog.jsp?page=clones/collaboration/index>

Seq primer: 5'-TTTTTTTTTTTTTTTTTTT(A/C/G)-3'.

Location/Qualifiers

1..773

/organism="Glycine max"

/mol\_type="mRNA"

/db\_xref="taxon:3847"

/clone="Gm-r1083-3531"

/clone\_lib="Gm-r1083"

/note="The library Gm-r1083 is a sequence-driven, reracked set of 4,992 clones selected from cDNA libraries from various tissues and stages of development of soybean. It represents 1117 sequences from the progenitor library Gm-cl009 (from mature roots of 2 month old greenhouse grown 'Williams' soybean plants); 820 sequences from the progenitor library Gm-cl013 (from 2 to 3 week old whole plants of Williams); and 3055 sequences from library Gm-cl028 (from 'Supernod' plants whose seedlings were inoculated with Bradyrhizobium japonicum, courtesy of Dr. Gary Stacey). The 5' ESTs of the source clones from the different progenitor libraries was used to select singletons, or a representative of each contig, which were reracked to form library Gm-r1083. The cDNA clones of the reracked Gm-r1083 library were then sequenced at the 3' end. The contig analysis to select unique genes was performed by the laboratory of Ernest Retzel, Center for Computational Genomics and Bioinformatics, University of Minnesota, <http://web.ahc.umn.edu/biodata/nsfsoy/>. Reracking was performed by Incyte Genomics, St. Louis, <http://www.incyte.com>, and 3' sequencing by the Keck Center for Comparative and Functional Genomics, University of Illinois, <http://www.lie.uiuc.edu/biotech/keck.html>. Note: The corresponding 5' EST from each clone in the Gm-r1083 library is listed in the 'OTHER EST' field. The detailed information on the source library for each clone can also be obtained by referring to the Incyte Genomics clone ID of the original cDNA library that is also listed under 'OTHER EST'."

ORIGIN

Alignment Scores:

Pred. No.: 0.00467 Length: 773

Score: 85.00 Matches: 16

Percent Similarity: 76.00% Conservative: 3

Best Local Similarity: 64.00% Mismatches: 6

Query Match: 57.82% Indels: 0

DB: 4 Gaps: 0

US-10-014-101B-32 (1-28) x BI970438 (1-773)

Qy 4 GlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSerAlaMet 23  
 |||||  
 684 GGACCACTCTGTGTTATCCATGATAGAAACAAGTGGTACGATGAGTGTGACGATCT 625  
 Db  
 Qy 24 IleProGluIleAsp 28  
 |||||  
 624 ATACGACGAGGAT 610  
 Db

RESULT 28  
 LOCUS CF209921  
 DEFINITION CAB20005\_IVb\_Fb\_G06 Cabernet Sauvignon Flower bloom EST 01-AUG-2003  
 vinifera cDNA clone CAB20005\_IVb\_Fb\_G06 5', mRNA sequence.  
 ACCESSION CF209921  
 VERSION CF209921.1 GI:33404294  
 KEYWORDS EST.  
 SOURCE Vitis vinifera  
 ORGANISM Vitis vinifera  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; Vitaceae; Vitis.  
 REFERENCE 1 (bases 1 to 809)  
 AUTHORS Goes da Silva,F., Iandolino,A., Lim,H., Baek,J., Leslie,A., Xu,J.,  
 Jones,K. and Cook,D.  
 TITLE Expressed sequence tags from Vitis vinifera 'Cabernet Sauvignon'  
 berries at various developmental stages  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: Douglas Cook, PhD  
 UC Davis, Plant Pathology  
 One Shields Ave, Davis, CA 95616, USA  
 Tel: 530 754 6561  
 Fax: 530 754 6617  
 Email: drcook@ucdavis.edu  
 Seq primer: ACGTACCGACATATGCC.

FEATURES  
 source  
 1..809  
 /organism="Vitis vinifera"  
 /mol\_type="mRNA"  
 /cultiVar="Cabernet Sauvignon"  
 /db\_xref="taxon:29760"  
 /clone="CAB20005\_IVb\_Fb\_G06"  
 /sex="Hermaphrodite"  
 /dev\_stage="Bloom"  
 /clone\_lib="Cabernet Sauvignon Flower bloom - CAB2"  
 /note="Organ: Flower - Bloom; Vector: pDNR; Site 1: SfiI;  
 Site 2: SfiI; CAB2 is a cDNA library of Vitis vinifera cv.  
 'Cabernet Sauvignon' Clone 8 berries. Samples were  
 collected at full bloom (80 to 100% flowers showing  
 dehiscence of calypters or caps and anthers fully  
 extended). Sampled vines were located at the University of  
 California, Davis, Experimental Vineyard. cDNAs were made  
 by oligo-dT priming and directionally cloned. 5' and 3'  
 adaptors were used in cloning as follows:  
 5'-AACGCTGATCAGCCAGAGTGGCCATTCGCGCGG-3' and  
 5'-ATTCTAGGCTGACGAGCGGCGACATG-dT(30)NN-3'. Library was  
 constructed using the Clontech Creator SMART kit and  
 size-selected to contain the 0.5-3 kb size fraction."

## ORIGIN

Alignment Scores:  
 Pred. No.: 0.00494 Length: 809  
 Score: 85.00 Matches: 15  
 Percent Similarity: 73.08% Conservative: 4  
 Best Local Similarity: 57.69% Mismatches: 7  
 Query Match: 57.82% Indels: 0  
 DB: 7 Gaps: 0

US-10-014-101B-32 (1-28) x CF209921 (1-809)

Qy 3 SerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22

Db 404 AGTGGCCCATCTCTATCTACCTATGACAAAAACAAGTGGGACGACCGCATCGTG 463  
 |||||  
 Qy 23 MetIleProGluIleAsp 28  
 |||||  
 464 GTTACGCCAGAGGAGGAC 481  
 Db

RESULT 29  
 LOCUS CL964867  
 DEFINITION OsIFCC011216 Oryza sativa Express Library Oryza sativa (indica  
 cultivar-group) genomic, genomic survey sequence.  
 ACCESSION CL964867  
 VERSION CL964867.1 GI:52384422  
 KEYWORDS GSS.  
 SOURCE Oryza sativa (indica cultivar-group)  
 ORGANISM Oryza sativa (indica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1 (bases 1 to 1587)  
 AUTHORS Ma,L., Wang,C.J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M.,  
 Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L.,  
 Wong,G.K.S., Deng,X.W. and Wang,J.  
 TITLE An analysis of transcriptional regulation of the rice genome and  
 its comparison to Arabidopsis  
 JOURNAL Unpublished (2004)  
 COMMENT Contact: Chen Chen  
 Department of Bioinformatic  
 Beijing Institute of Genomics  
 Chinese Academy of Sciences, Beijing 101300, China  
 Tel: 86-10-80481559  
 Fax: 86-10-80488676  
 Email: chenchen@genomics.org.cn  
 Rice genomic sequence.  
 Class: exon-trapped.

FEATURES  
 source  
 1..1587  
 Location/Qualifiers  
 /organism="Oryza sativa (indica cultivar-group)"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:39946"  
 /clone\_lib="Oryza sativa Express Library"  
 /note="Oryza sativa exon trapped genomic sequences"

## ORIGIN

Alignment Scores:  
 Pred. No.: 0.0113 Length: 1587  
 Score: 85.00 Matches: 14  
 Percent Similarity: 80.00% Conservative: 6  
 Best Local Similarity: 56.00% Mismatches: 5  
 Query Match: 57.82% Indels: 0  
 DB: 9 Gaps: 0

US-10-014-101B-32 (1-28) x CL964867 (1-1587)

Qy 4 GlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSerAlaMet 23  
 |||||  
 Db 1255 GGCCTTCATCTCATGTACCCCATGATGAATAGGACATGTGGATGACAGGATGACGCGCATG 1314  
 |||||  
 Qy 24 IleProGluIleAsp 28  
 |||||  
 Db 1315 ACGCGGAGGAGGAC 1329

## ORIGIN

Alignment Scores:  
 Pred. No.: 0.00494 Length: 809  
 Score: 85.00 Matches: 15  
 Percent Similarity: 73.08% Conservative: 4  
 Best Local Similarity: 57.69% Mismatches: 7  
 Query Match: 57.82% Indels: 0  
 DB: 7 Gaps: 0

US-10-014-101B-32 (1-28) x CF209921 (1-809)

Qy 3 SerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22



```

RESULT 32
AV536711
LOCUS AV536711 481 bp mRNA linear EST 20-FEB-2004
DEFINITION Arabidopsis thaliana liquid-cultured seedlings Columbia
Arabidopsis thaliana cDNA clone pAZNI10455R 5', mRNA sequence.
ACCESSION AV536711
VERSION AV536711.1 GI:8696994
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 481)
Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
DNA Res. 7 (3), 175-180 (2000)
JOURNAL 20363093
MEDLINE 10907847
PUBMED
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
FEATURES
source
1..481
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="pAZNI10455R"
/tissue_type="liquid-cultured seedlings"
/clone_lib="Arabidopsis thaliana liquid-cultured seedlings
Columbia"
/notes="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"
ORIGIN
Alignment Scores:
Pred. No.: 0.00549 Length: 481
Score: 83.00 Matches: 19
Percent Similarity: 37.93% Conservative: 3
Best Local Similarity: 32.76% Mismatches: 6
Query Match: 56.46% Indels: 30
DB: 1 Gaps: 1
US-10-014-101b-32 (1-28) x AV536711 (1-481)
QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnlys----- 15
Db 209 TCAACTTCGTGTTACTCTCTCTATCCACAAACCGAACAAGTAATATTACTTTT 268
QY 15 ----- 15
Db 269 TGATTTTGTTTTATTTGAAGTATATCCCAATAATGTATGTAAATTTGTTAAACAAGATT 328
QY 16 -----TtpAspAsnArgMetSerAlaMetIleProGluIleAsp 28
Db 329 TATTATTAAATAGATGGACACCCGATGTCTACCATGACACCGGACGAAGAT 382
RESULT 33
CL947867
LOCUS CL947867 1566 bp DNA linear GSS 21-SEP-2004
DEFINITION OsIFSB001488 Oryza sativa Express Library Oryza sativa (indica
cultivar-group) genomic, genomic survey sequence.
ACCESSION CL947867
VERSION CL947867.1 GI:52359876
KEYWORDS GSS.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)

```

```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 1566)
Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,
Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
Wong, G.K.S., Deng, X.W. and Wang, J.
An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
Unpublished (2004)
JOURNAL Contact: Chen Chen
COMMENT Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.
FEATURES
source
1..1566
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Express Library"
/notes="Oryza sativa exon trapped genomic sequences"
ORIGIN
Alignment Scores:
Pred. No.: 0.0234 Length: 1566
Score: 83.00 Matches: 14
Percent Similarity: 76.92% Conservative: 6
Best Local Similarity: 53.85% Mismatches: 6
Query Match: 56.46% Indels: 0
DB: 9 Gaps: 0
US-10-014-101b-32 (1-28) x CL947867 (1-1566)
QY 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnlysTrpAspAsnArgMet 20
Db 1201 AGTAACAATGTCCTATCTCTATCCAGTGACAAATCCAGATGGGACACAGAAACA 1260
QY 21 SerAlaMetIleProGlu 26
Db 1261 TCAGTAGTCATACACAGAT 1278
RESULT 34
BF520835
LOCUS BF520835 575 bp mRNA linear EST 27-JAN-2003
DEFINITION EST458308 DSIL Medicago truncatula cDNA clone pDSL-39F7, mRNA
sequence.
ACCESSION BF520835
VERSION BF520835.1 GI:11609518
KEYWORDS EST.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
1 (bases 1 to 575)
Fedorova, M., Pierson, B.L., Samac, D.A., Vance, C.P., Gantt, G.S.,
Peng, H., Ellis, L., Town, C.D., Bowman, C.L., Craven, M.B.,
Hansen, T.S., Holt, I.E. and Fraser, C.M.
ESTs from leaves of Medicago truncatula after inoculation with
Colletotrichum trifolii
Unpublished (2000)
JOURNAL Contact: Deborah A. Samac
COMMENT Department of Plant Pathology
University of Minnesota
495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
Tel: 612 625 1243
Fax: 651 649 5058

```



```

/clone_lib="MTUS"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
was directionally ligated into the Unizap XR vector from
Stratagene and packaged using Gigapack III Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant lambda-Zap phage using Ex-assist
helper phage and propagated in XL0LR cells."

ORIGIN

Alignment Scores:
Pred. No.: 0.0112 Length: 636
Score: 82.00 Matches: 13
Percent Similarity: 75.00% Conservative: 8
Best Local Similarity: 46.43% Mismatches: 7
Query Match: 55.78% Indels: 0
DB: 6 Gaps: 0

US-10-014-101B-32 (1-28) x CA918993 (1-636)

Qy 1 SerAlaSerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLyTrpAspAsnArgMet 20
Db 528 ACAAGCAATGCCCTGCTTATCTATCCAGTGAAATAATCAAGTGGGACATAGAACT 469

Qy 21 SerAlaMetIleProGluIleAsp 28
Db 468 TCAGTTGTTATCCAGATGAAGAT 445

RESULT 37
LOCUS CC693605 753 bp DNA linear GSS 19-JUN-2003
DEFINITION OGULX22TH_ZM_0.7_1.5_KB_Zea_mays_genomic_clone_ZMBM0471C20,
genomic survey sequence.
ACCESSION CC693605
VERSION CC693605.1 GI:32098381
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 753)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other GSSs: OGULX22TV
Contact: Cathy Whitelaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source Location/Qualifiers
1..753
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMBM0471C20"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Alignment Scores:
Pred. No.: 0.0138 Length: 753
Score: 82.00 Matches: 12
Percent Similarity: 86.96% Conservative: 8
Best Local Similarity: 52.17% Mismatches: 3

/clone_lib="MTUS"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
was directionally ligated into the Unizap XR vector from
Stratagene and packaged using Gigapack III Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant lambda-Zap phage using Ex-assist
helper phage and propagated in XL0LR cells."

ORIGIN

Alignment Scores:
Pred. No.: 0.0146 Length: 788
Score: 82.00 Matches: 13
Percent Similarity: 75.00% Conservative: 8
Best Local Similarity: 46.43% Mismatches: 7
Query Match: 55.78% Indels: 0
DB: 6 Gaps: 0

US-10-014-101B-32 (1-28) x CF068302 (1-788)

Query Match: 55.78% Indels: 0
DB: 9 Gaps: 0

US-10-014-101B-32 (1-28) x CC693605 (1-753)

Qy 3 SerGlyLeuAlaLeuLeuTyProThrAsnArgAsnLyTrpAspAsnArgMetSerAla 22
Db 205 GCTGGGATCATCTCTATGATACCCATGAACAGGACAGGTGGACGCCGATGACAGCG 264

Qy 23 MetIlePro 25
Db 265 ATGACCCCA 273

RESULT 38
LOCUS CF068302 788 bp mRNA linear EST 21-JUL-2003
DEFINITION EST669023 MTUS Medicago truncatula cDNA clone MTUS-6C8, mRNA
sequence.
ACCESSION CF068302
VERSION CF068302.1 GI:33104961
KEYWORDS EST.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 788)
AUTHORS VandenBosch,K., Endre,G., Silverstein,K, Town,C.D., Van Aken,S.,
Utterback,T., Cheung,F. and Fraser,C.M.
TITLE The Medicago truncatula 6K unigene set: cDNA clones selected and
re-arranged from various libraries
JOURNAL Unpublished (2002)
COMMENT Contact: VandenBosch K
Department of Plant Biology
University of Minnesota
220 BioSci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
Tel: 612 624 2755
Fax: 612 625 1738
Email: kvandenb@bcs.umn.edu
TIGR sequence name: MTUSF32TK More information is available at:
www.medicago.org
Seq primer: SKmod (CTA GAA CTA gTg GAT CC).
Location/Qualifiers
1..788
/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="Al7"
/db_xref="taxon:3880"
/clone="MTUS-6C8"
/tissue_type="mixed tissues"
/dev_stage="various stages"
/lab_host="XL0LR"
/clone_lib="MTUS"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
was directionally ligated into the Unizap XR vector from
Stratagene and packaged using Gigapack III Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant lambda-Zap phage using Ex-assist
helper phage and propagated in XL0LR cells."

```

Qy 1 SerAlaSerGlyLeuAlaLeuLeuTyPProThrAsnArgAsnLysTrpAspAsnArgMet 20  
 Db ACAGCAATGCGCCCTGTCCTTATCTTATCCAGTCGAAAGAAATCAAAGTGGGCAATAGAACT 367  
 Qy 21 SerAlaMetIleProGluIleAsp 28  
 Db TCAGTTGTTATTCAGATGAAGAT 391

RESULT 39  
 CO483046/c

LOCUS CO483046 831 bp mRNA linear EST 09-JUL-2004

DEFINITION GQ0198.TB A09 GQ019: Clean ROOTS systems - P treatments Picea glauca cDNA clone GQ0198\_A09 3', mRNA sequence.

ACCESSION CO483046

VERSION CO483046.1 GI:50164789

SOURCE EST.

ORGANISM Picea glauca (white spruce)

REFERENCE Picea glauca

AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea. 1 (bases 1 to 831)

TITLE Morency M., J., Cooke J., Pavy N., Parsons L., Paule C., Seguin A., Retzel E., Butterfield Y., Barber S., Yang G., Scott J., Siddiqui A., Holt R., Marra M. and Mackay J.

JOURNAL Arborea EST sequencing in Picea glauca (white spruce) Unpublished (2004)

COMMENT Contact: John Mackay  
 Centre de Recherche en Biologie Forestiere  
 Universite Laval  
 Pavillon Charles-Eugene Marchand, Quebec, Canada G1K 7P4  
 Fax: 418 656 7493  
 Email: jmackay@rsvs.ulaval.ca  
 Center for Computational Genomics and Bioinformatics (CCGB),  
 University of Minnesota, MN id Identifier: MN5256147 Clone ID:  
 GQ0198\_A09 Clones available through: John Mackay, Ph. D. Professeur  
 adjoint - Assistant professeur EMAIL: jmackay@rsvs.ulaval.ca Centre  
 de Recherche en Biologie Forestiere (Forest Biology Research  
 Center) Universite Laval Quebec, Quebec Canada G1K 7P4  
 Plate: 8 row: 09 column: A  
 Seq primer: PolyT plus Primer.  
 Location/Qualifiers  
 1. 831  
 /organism="Picea glauca"  
 /mol\_type="mRNA"  
 /strain="pg-653"  
 /db\_xref="taxon:3330"  
 /clones="GQ0198\_A09"  
 /sex="Hermaphrodite"  
 /tissue\_type="Whole roots systems, predominantly free of  
 michorizes"  
 /lab\_host="E. coli DH10B cells"  
 /clone\_lib="GQ019: Clean ROOTS systems - P treatments"  
 /note="Organ: Whole root systems form 15 cm tall, 4 month  
 old seedlings produced by somatic embryogenesis; Vector:  
 pBluescript II SK (-) XR; Site 1: Eco-RI; Site 2: Xho-I;  
 Tissues were pooled from several plants receiving daily  
 fertilization varying potassium levels (either 0.02 mM or 2  
 mM PO4), collected at 4 hours, 24 hours, 2 days, 4 days or 7  
 days of treatment. cDNA was prepared from 5 mg of poly  
 A+ selected RNA and was directionally ligated into the  
 pBluescript II SK (-) XR vector (Stratagene), transformed  
 by electroporation into DH10B cells (in vitro) for  
 propagation"

ORIGIN

Alignment Scores:  
 Pred. No.: 0.0156 Length: 831  
 Score: 82.00 Matches: 15  
 Percent Similarity: 76.92% Conservative: 5  
 Best Local Similarity: 57.69% Mismatches: 6  
 Query Match: 55.78% Indels: 0  
 Gaps: 7

DB: 7 Gaps: 0  
US-10-014-101B-32 (1-28) x CO485966 (1-856)  
Qy 3 SerGlyLeuAlaLeuLeuTyrProThrAsnArgAsnLysTrpAspAsnArgMetSerAla 22  
Db 406 AATGGATTAAAGCTCTGTCCTCTGAGAGAAAGCTTGTGGATTCTCGACGTCGGCT 347  
Qy 23 MetIleProGluIleAsp 28  
Db 346 GTTATCCAGAGGAGGAT 329

Search completed: February 18, 2005, 14:07:05  
Job time : 7208 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 17, 2005, 21:59:48 ; Search time 959.84 Seconds  
(without alignments)  
353.378 Million cell updates/sec

Title: US-10-014-101B-39

Perfect score: 35

Sequence: 1 VGGTSLN 7

Scoring table:

BLOSUM62	Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2 1/USPTO.spool/US10014101/runat 16022005 075845 7936/app query.fasta\_1.796  
-DB=GenEmbl -OPMT=fastap -SUFFIX=rge -MINMATCH=0 -LOOPTCL=0 -LOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=40 -MODE=LOCAL  
-OUTFMT=ptc -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10014101 @CGN 1 1 2886 @runat 16022005 075845 7936 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	35	100.0	116	11	BX295789 Arabidops
c 2	35	100.0	379	8	AY444352 Pisum sat
3	35	100.0	424	11	BX323242 Arabidops
c 4	35	100.0	442	6	AX886741 Sequence

C	5	6	BD026351	442	6	BD026351	Sequence
35	100.0	1506	CQ812637	1506	6	CQ812637	Sequence
35	100.0	1506	AX339729	1506	6	AX339729	Sequence
35	100.0	1515	AX785077	1515	6	AX785077	Sequence
35	100.0	1515	AX339733	1515	6	AX339733	Sequence
35	100.0	1515	AX785081	1515	6	AX785081	Sequence
35	100.0	1537	BT005653	1537	8	BT005653	Arabidops
35	100.0	1566	AX654568	1566	6	AX654568	Sequence
35	100.0	1572	AX339730	1572	6	AX339730	Sequence
35	100.0	1572	AX785078	1572	6	AX785078	Sequence
35	100.0	1572	AF303979	1572	8	AF303979	Arabidops
35	100.0	1575	AX339731	1575	6	AX339731	Sequence
35	100.0	1575	AX507394	1575	6	AX507394	Sequence
35	100.0	1575	AX651615	1575	6	AX651615	Sequence
35	100.0	1575	AX785079	1575	6	AX785079	Sequence
35	100.0	1575	AF303980	1575	8	AF303980	Arabidops
35	100.0	1587	AF303981	1587	6	AF303981	Arabidops
35	100.0	1590	AX653027	1590	6	AX653027	Sequence
35	100.0	1590	AX653226	1590	6	AX653226	Sequence
35	100.0	1593	AX652883	1593	6	AX652883	Sequence
35	100.0	1605	AR150952	1605	6	AR150952	Sequence
35	100.0	1605	AR399211	1605	6	AR399211	Sequence
35	100.0	1606	BT002757	1606	8	BT002757	Arabidops
35	100.0	1608	AX039923	1608	6	AX039923	Sequence
35	100.0	1611	AX339732	1611	6	AX339732	Sequence
35	100.0	1611	AX785080	1611	6	AX785080	Sequence
35	100.0	1620	AX339737	1620	6	AX339737	Sequence
35	100.0	1620	AX785085	1620	6	AX785085	Sequence
35	100.0	1623	AF303982	1623	8	AF303982	Arabidops
35	100.0	1655	BT000179	1655	8	BT000179	Arabidops
35	100.0	1677	AX653630	1677	6	AX653630	Sequence
35	100.0	1687	BT004107	1687	8	BT004107	Arabidops
35	100.0	1728	CQ812635	1728	6	CQ812635	Sequence
35	100.0	1728	AX339728	1728	6	AX339728	Sequence
35	100.0	1728	AX785076	1728	6	AX785076	Sequence
35	100.0	1776	ZMW18377	1776	8	ZMW18377	Ze mays mr
35	100.0	1857	AF540382	1857	8	AF540382	Hordeum v
35	100.0	1863	AY054460	1863	8	AY054460	Arabidops
35	100.0	1866	AY0209184	1866	8	AY0209184	Hordeum v
35	100.0	1873	AY091158	1873	8	AY091158	Arabidops
35	100.0	1930	AK101022	1930	8	AK101022	Oryza sat

#### ALIGNMENTS

RESULT 1  
BX295789  
LOCUS  
DEFINITION  
Arabidopsis thaliana transposon insertion STS SM\_3.33583, sequence tagged site.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
BX295789.1 GI:29169627  
STs: STS, sequence tagged site.  
Arabidopsis thaliana (thale cress)

REFERENCE  
AUTHORS  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
1  
Clarke, J.H., Bowles, B., Carter, J., Hart, D., McCullagh, B., Murphy, G., Langham, S., Legrys, C., Jones, J.D.G. and Bevan, M. Unpublished  
2 (bases 1 to 116)  
Clarke, J.H.  
Submitted (24-WAR-2003) Clarke J.H., John Innes Centre, Colney Lane, Norwich, NR4 7UJ, UK  
Ar denotes an activation tag dissociation transposon within a single line, ET an enhancer trap dissociation transposon, GT a gene trap dissociation transposon, MT a mis-expression enhancer trap dissociation transposon, SM a defective suppressor mutator transposon. \_3 denotes a sequence derived from the 3' end of the transposon. \_5 denotes a sequence derived from the 5' end of the

transposon BBSRC GARNET, ATIS project  
On-line seed stock requests: [http://nasc.nott.ac.uk/NASC\\_stock\\_code/N120294](http://nasc.nott.ac.uk/NASC_stock_code/N120294).

## FEATURES

```

1. . . . .116
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variety="Columbia-0 NASC stock code N1092"
/db_xref="taxon:3702"
/clone="AL079344"
/note="Derived from superpool 1.32 NASC code N40031"
1. . . . .116
/standard_name="SM 3.33583"

```

## STS

```

/NOTE= DERIVED FROM SUPERPOOL 1.152 AND CC CODE 110000
1. .116
/standard name="SM 3.33583"

```

## ORIGIN

Alignment Scores:		
Pred. No.:	78.9	Length: 116
Score:	35.00	Matches: 7
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	11	Gaps: 0

US-10-014-101B-39 (1-7) x BX295789 (1-116)

Qy 1 ValGlyGlyThrLeuSerAsn 7  
Db 80 GTGGCGGGACGTTGTCTGAAC 10

## RESULT 2

AY444352/c	AY444352	379 bp	DNA linear	PLN 16-DEC-2003
LOCUS	Pisum sativum cytokinin oxidase/dehydrogenase 1 (PCRX1) gene,			
DEFINITION	partial sequence.			

ACCESSION  
AY444352

ACCESSION	AY444352	
VERSION	AY444352.1	GI:39636914
KEYWORDS		
SOURCE	<i>Pisum sativum</i> (pea)	
ORGANISM	<i>Pisum sativum</i>	
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; <i>Pisum</i> .	

REFERENCE 1 (bases 1 to 379)

**AUTHORS** Vaseva-Gemiteva, I.I., Lee, D., Turner, L., Ellis, N. and Karanov, E.N.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (21-OCT-2003) Plant Physiology Institute, Bulgarian Academy of Sciences, Georgi Bonchev 21 Building, Sofia 1113, Bulgaria

## FEATURES

```

FEATURES
  source
    location/Qualifiers
      1. .379
        /organism="Pisum sativum"
        /mol_type="genomic DNA"
        /cultivar="Manuela"
        /db_xref="taxon:3888"
        /chromosome="6"
        /tissue_type="leaves"
      <1. .>379
        /gene="PCKX1"

```

gene

gene  
1. 2373  
/gene="PCKX1"  
/note="cytokinin oxidase/dehydrogenase 1"

**ORIGIN**

Alignment Scores:	214	Length:	379
Pred. No.:		Matches:	7
Score:	35.00	Conservative:	0
Percent Similarity:	100.0%	Mismatches:	0
Best Local Similarity:	100.0%	Indels:	0
Query Match:	100.0%	Gaps:	0
DB:	8		

US-10-014-101B-39 (1-7) x AY444352 (1-379)

1 ValGlyGlyThrLeuSerAsn 7 Qv

Db 160 GTGGGTGGAACGCTCTCTAAC 140

### RESULT 3

Accession	EX323242	424 bp	DNA	linear	STS 10-JUN-2003
LOCUS	EX323242				
DEFINITION	Arabidopsis thaliana transposon insertion STS SM_3.38194, sequence tagged site.				
ACCESSION	EX323242				
VERSION	EX323242.1				GI:29602993
KEYWORDS	STS; STS, sequence tagged site.				
SOURCE	Arabidopsis thaliana (thale cress)				

## REFERENCE

CLARKE, J. H., BOWLES, B., CARTER, J., HART, D., MCCULLAGH, B.,  
MURPHY, G., LANGHAM, S., LEGRYS, C., JONES, J. D. G. and BEVAN, M.  
JOURNAL  
REFERENCE  
2 (bases 1 to 424)  
AUTHORS  
CLARKE, J. H.

**TITLE**  
**Direct Submission**

**JOURNAL**  
Submitted (07-APR-2003) Clarke J.H., John Innes Centre, Colney Lane, Norwich, NR4 7UJ, UK  
**COMMENT**  
AT denotes an activation tag dissociation transposon within a single line, ET an enhancer trap dissociation transposon, GT a trap dissociation transposon, MI a mis-expression enhancer trap dissociation transposon, SM a defective suppressor mutator transposon. \_3 denotes a sequence derived from the 3' end of the transposon, \_5 denotes a sequence derived from the 5' end of the transposon. BBSRC GARNET, ARIS project  
On-line seed stock requests: <http://nasc.nott.ac.uk/> NASC stock code: N124905.

## FEATURES

```

1. .424
source
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variety="Columbia-0" NASC stock code N1092"
/db_xref="taxon:3702"
/clone="AC023754"
/note="Derived from superpool 15.46 NASC collection"
1. .424
STS
/standard_name="SM 3.38194"

```

## STS

ORIGIN				
Alignment Scores:				
Pred. No.:	235	Length:		424
Score:	35.00	Matches:		7
Percent Similarity:	100.00%	Conservative:		0
Best Local Similarity:	100.00%	Mismatches:		0
Query Match:	100.00%	Indels:		0
DR:	11	Gaps:		0

DB: II GA  
US-10-014-101B-39 (1-7) x BX323242 (1-424)

Qv 1 ValGlvGlvThrlEiSerAsn 7

QY 1 valgiygiyihlleuseelash /  
|||||  
365 GTGGAGGTACACTCTCCAAT 385  
Db

## RESULT 4

AX886741.1	AX886741/c	AX886741	442 bp	DNA	linear	PAT 18-DEC-2003
LOCUS		Sequence	2604	from Patent EP1033401.		
DEFINITION		AX886741				
ACCESSION		AX886741				
VERSION		AX886741.1	GI:40004205			

## KEYWORDS

1  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE Dumas Milne Edwards, J.B., Duclert, A. and Giordano, J.Y.

**TITLE** Expressed sequence tags and encoded human proteins  
**JOURNAL** Patent: EP 1033401-A 2604 06-SEP-2000;  
**GENSET** (FR)  
**FEATURES** Location/Qualifiers  
    **source**  
        1..442  
        /organism="Homo sapiens"  
        /mol\_type="unassigned DNA"  
        /db\_xref="taxon:9606"  
        126..>440  
        /notes="unnamed protein product"  
        /codon\_start=1  
        /protein\_id="CAF00879.1"  
        /db\_xref="GI:40044206"  
        /translation="MQAQSKEQQLDINKYQQLSRLDEILSRIAKETBEIKDLE  
        EOLTEGQIAANEALKKDLGVISGLQYLGTIKGQATQAQNECKLRDKETLLQRLT  
        EVE"  
**CDS**  
        126..>440  
        /organism="Homo sapiens"  
        /mol\_type="genomic DNA"  
        /db\_xref="taxon:9606"  
**ORIGIN**  
    Alignment Scores:  
    Pred. No.: 243 Length: 442  
    Score: 35.00 Matches: 7  
    Percent Similarity: 100.00% Conservative: 0  
    Best Local Similarity: 100.00% Mismatches: 0  
    Query Match: 100.00% Indels: 0  
    DB: 6 Gaps: 0  
    US-10-014-101B-39 (1-7) x AX886741 (1-442)  
**QY** 1 ValGlyGlyThrLeuSerAen 7  
    |||||  
**Db** 67 GTGGCGGGAACGCTTCTTAAT 47  
**RESULT 5**  
**LOCUS** BD026351/c 442 bp DNA linear PAT 27-AUG-2002  
**DEFINITION** Sequence tag and encoded human protein.  
**ACCESSION** BD026351  
**VERSION** BD026351.1 GI:22567574  
**KEYWORDS** JP 2001269182-A/2597.  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens  
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
**REFERENCE**  
**AUTHORS** Edwards, J.B.D.M., Duclair, E. and Jordan, J.Y.  
**TITLE** Sequence tag and encoded human protein  
**JOURNAL** Patent: JP 2001269182-A 2597 02-OCT-2001;  
**GENSET**  
**COMMENT**  
    OS Homo sapiens (human)  
    PN JP 2001269182-A/2597  
    PD 02-OCT-2001  
    PF 24-FEB-2000 JP 2000118773  
    PR 26-FEB-1999 US 60/122487  
    PI JEAN BAPTISTE DUMAS MILNE EDWARDS, EIMERIC DUCLAIR, JEAN YVES  
    PI JORDAN  
    PC C12N15/09, C07K14/435, C07K16/18, C12N1/15, C12N1/19, C12N1/21, PC  
    C12N5/10,  
    PC C12P21/02, C12P21/08, C12Q1/68//G06F17/30, C12N15/00, C12N5/00, PC  
    G06F15/40  
**CC** G06F15/40  
**FT** Key Location/Qualifiers  
    CDS 126..440.  
**FEATURES** Location/Qualifiers  
    **source**  
        1..442  
        /organism="Homo sapiens"  
        /mol\_type="genomic DNA"  
        /db\_xref="taxon:9606"  
**ORIGIN**  
    Alignment Scores:  
    Pred. No.: 243 Length: 442  
    Score: 35.00 Matches: 7  
    Percent Similarity: 100.00% Conservative: 0  
**Best Local Similarity:** 100.00%  
**Query Match:** 100.00%  
**Mismatches:** 0  
**Indels:** 0  
**Gaps:** 0

---

**Best Local Similarity:** 100.00%  
**Query Match:** 100.00%  
**Mismatches:** 0  
**Indels:** 0  
**Gaps:** 0  
**US-10-014-101B-39 (1-7) x BD026351 (1-442)**  
**QY** 1 ValGlyGlyThrLeuSerAen 7  
    |||||  
**Db** 67 GTGGCGGGAACGCTTCTTAAT 47  
**RESULT 6**  
**LOCUS** CQ812637 1506 bp DNA linear PAT 24-MAY-2004  
**DEFINITION** Sequence 3 from Patent WO2004038027.  
**ACCESSION** CQ812637  
**VERSION** CQ812637.1 GI:47602086  
**KEYWORDS** Arabidopsis thaliana (thale cress)  
**SOURCE** Arabidopsis thaliana  
**ORGANISM** Arabidopsis thaliana  
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
    rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsia.  
**REFERENCE**  
**AUTHORS** van Camp, W.  
**TITLE** Bioremediation  
**JOURNAL** Patent: WO 2004038027-A 3 06-MAY-2004;  
**CropDesign** N.V. (BE)  
**FEATURES** Location/Qualifiers  
    **source**  
        1..1506  
        /organism="Arabidopsis thaliana"  
        /mol\_type="unassigned DNA"  
        /db\_xref="taxon:3702"  
**ORIGIN**  
    Alignment Scores:  
    Pred. No.: 682 Length: 1506  
    Score: 35.00 Matches: 7  
    Percent Similarity: 100.00% Conservative: 0  
    Best Local Similarity: 100.00% Mismatches: 0  
    Query Match: 100.00% Indels: 0  
    DB: 6 Gaps: 0  
**US-10-014-101B-39 (1-7) x CQ812637 (1-1506)**  
**QY** 1 ValGlyGlyThrLeuSerAen 7  
    |||||  
**Db** 466 GTGGGAGGACGTTGCGAAT 486  
**RESULT 7**  
**LOCUS** AX339729 1506 bp DNA linear PAT 10-JAN-2002  
**DEFINITION** Sequence 26 from Patent WO0196580.  
**ACCESSION** AX339729  
**VERSION** AX339729.1 GI:18135722  
**KEYWORDS** Arabidopsis thaliana (thale cress)  
**SOURCE** Arabidopsis thaliana  
**ORGANISM** Arabidopsis thaliana  
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
    rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsia.  
**REFERENCE**  
**AUTHORS** Schmuelling, T. and Werner, T.  
**TITLE** Method for modifying plant morphology, biochemistry and physiology  
**JOURNAL** Patent: WO 0196580-A 26 20-DEC-2001;  
**Schmuelling, Thomas (DE) ; Werner, Tomas (DE)**  
**FEATURES** Location/Qualifiers  
    **source**  
        1..1506  
        /organism="Arabidopsis thaliana"  
        /mol\_type="unassigned DNA"  
        /db\_xref="taxon:3702"  
**ORIGIN**  
    Alignment Scores:  
    Pred. No.: 682 Length: 1506  
    Score: 35.00 Matches: 7  
    Percent Similarity: 100.00% Conservative: 0  
    Best Local Similarity: 100.00% Mismatches: 0  
    Query Match: 100.00% Indels: 0  
    DB: 6 Gaps: 0

```

Pred. No.: 682          Length: 1506
Score: 35.00           Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6                  Gaps: 0

US-10-014-101B-39 (1-7) x AX339729 (1-1506)

QY 1 ValGlyGlyThrLeuSerAsn 7
Db 466 GTCGGAGGACGTTGTCCAAAT 486

RESULT 8
AX785077 1506 bp DNA linear PAT 17-JUL-2003
DEFINITION Sequence 26 from Patent WO03050287.
ACCESSION AX785077
VERSION AX785077.1 GI:32952908
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
TITLE rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
JOURNAL
FEATURES
source
1. 1506
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Alignment Scores:
Pred. No.: 682          Length: 1506
Score: 35.00           Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6                  Gaps: 0

US-10-014-101B-39 (1-7) x AX785077 (1-1506)

QY 1 ValGlyGlyThrLeuSerAsn 7
Db 466 GTCGGAGGACGTTGTCCAAAT 486

RESULT 9
AX339733 1515 bp DNA linear PAT 10-JAN-2002
DEFINITION Sequence 30 from Patent WO0196580.
ACCESSION AX339733
VERSION AX339733.1 GI:18135726
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
TITLE rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
JOURNAL
FEATURES
source
1. 1515
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Alignment Scores:
Pred. No.: 682          Length: 1506
Score: 35.00           Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6                  Gaps: 0

US-10-014-101B-39 (1-7) x AX785077 (1-1506)

QY 1 ValGlyGlyThrLeuSerAsn 7
Db 466 GTCGGAGGACGTTGTCCAAAT 486

RESULT 10
AX785081 1515 bp DNA linear PAT 17-JUL-2003
DEFINITION Sequence 30 from Patent WO03050287.
ACCESSION AX785081
VERSION AX785081.1 GI:32952912
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
TITLE rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
JOURNAL
FEATURES
source
1. 1515
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Alignment Scores:
Pred. No.: 685          Length: 1515
Score: 35.00           Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6                  Gaps: 0

US-10-014-101B-39 (1-7) x AX339733 (1-1515)

QY 1 ValGlyGlyThrLeuSerAsn 7
Db 493 GTAGGTGGTACTCTGTCCAAAT 513

RESULT 11
BT005653 1537 bp mRNA linear PLN 15-MAR-2003
DEFINITION Arabidopsis thaliana clone U20989 putative cytokinin oxidase
ACCESSION BT005653
VERSION BT005653.1 GI:28973496
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
TITLE rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
JOURNAL
FEATURES
source
1 (bases 1 to 1537)
Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
Onodera,C.S., Quach,H.L., Tang,C., Toriumi,M., Wong,C., Wu,H.C.,
Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y.,
Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M.,
Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P.,

```

Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W.,  
Ecker, J.R. and Theologis, A.  
Arabidopsis Open Reading Frame (ORF) Clones  
Unpublished  
2 (bases 1 to 1537)  
Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M.,  
Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C.,  
Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y.,  
Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M.,  
Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P.,  
Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W.,  
Ecker, J.R. and Theologis, A.  
Direct Submission  
Submitted (15-MAR-2003) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA  
The RIKEN Genomic Sciences Center (GSC) members carried out the  
collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN  
Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J.,  
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,  
Hayashizaki, Y., and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members constructed and  
sequenced the PUNI (ORF) clones using the RAPL cDNAs: Yamada, K.,  
Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M.,  
Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C.,  
Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M.,  
Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W.,  
Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally  
to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP  
/PGEC) contributed equally to this work as PIs.

Annotation based on July 2002 version of the Arabidopsis genome  
submitted to Genbank.

**FEATURES**  
source  
1..1537  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="2"  
/clone="U20989"  
/ecotype="Columbia"  
/notes="This clone is in pUNI 51."  
1..1537  
/gene="At2g19500"  
1..1506  
/gene="At2g19500"  
/codon\_start=1  
/evidence=experimental  
/product="putative cytokinin oxidase"  
/protein\_id="AA064073.1"  
/db\_xref="GI:28973497"  
translations="MANLRIMLTITLVMTKSSNGIKIDLPKSLNLTSTDPISLSA  
ASDFGNITVTGGVCKPSSTADISRLQYANGSTQFQVARGGSHNGQASVSG  
GVIVNMTCTIDVVSKDCKYADVAAAGTLVVDLTKTAKGSPVSWTDYLIHTVGGTL  
SNGGICQVFRNGPLVSNVLELDVITCKGEMLTCSQLNPELFYGLGIGQFIITR  
ARTVLHAPKRAKFWRLMYSDFTTFTKQDERLSMANDIGVDYLEGQIFLSNGVVDTR  
FPPSQSKVADYVKQHGIIYVLEVAKYDDPNLPIISKVIDTTLTSLVLPQFISMH  
DVAYFPLARVHVEKNKLSLGLWELPHWNLNLYPKSRILDFHNGVVDKLLKQSA  
SGLALYPTNRNKNWDRMSMIFEDIEDVYIIGLQSATPKDLPEVSNKLIIRFC  
KDSGIIKQYLMHYTSKEDWIEHFGSKWDDFSKRKOLFDPKLLSPGQDIF"  
1507..1537  
/gene="At2g19500"

**3' UTR**  
ORIGIN  
Alignment Scores:  
Pred. No.: 694 Length: 1537  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0

**US-10-014-101B-39 (1-7) x BT005653 (1-1537)**  
QY 1 ValGlyGlyThrLeuSerAen 7  
Db 466 GTCGGAGGAACGTTGTGGAAT 486  
RESULT 12  
LOCUS AX654568  
DEFINITION Sequence 4438 from Patent WO03000898.  
ACCESSION AX654568  
VERSION AX654568.1 GI:29157382  
KEYWORDS Oryza sativa  
SOURCE Oryza sativa  
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
REFERENCE 1  
AUTHORS Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,  
Katagiri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.  
TITLE Plant genes involved in defense against pathogens  
JOURNAL Patent: WO 03000898-A 4438 03-JAN-2003;  
Syngenta Participations AG (CH)  
FEATURES  
source  
1..1566  
/organism="Oryza sativa"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4530"

**ORIGIN**  
Alignment Scores:  
Pred. No.: 705 Length: 1566  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

**US-10-014-101B-39 (1-7) x AX654568 (1-1566)**  
QY 1 ValGlyGlyThrLeuSerAen 7  
Db 499 GTTGGAGGACACTGTGGAAT 519  
RESULT 13  
LOCUS AX339730  
DEFINITION Sequence 27 from Patent WO0196580.  
ACCESSION AX339730  
VERSION AX339730.1 GI:18135723  
KEYWORDS Arabidopsis thaliana (thale cress)  
SOURCE Arabidopsis thaliana  
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi  
REFERENCE 1  
AUTHORS Schmullling, T. and Werner, T.  
TITLE Method for modifying plant morphology, biochemistry and physiology  
JOURNAL Patent: WO 0196580-A 27 20-DEC-2001;  
Schmullling, Thomas (DE) ; Werner, Tomas (DE)  
FEATURES  
source  
1..1572  
/organism="Arabidopsis thaliana"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3702"

**ORIGIN**  
Alignment Scores:  
Pred. No.: 707 Length: 1572  
Score: 35.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-10-014-101B-39 (1-7) x AX339730 (1-1572)

Qy 1 ValGlyGlyThrLeuSerAsn 7  
 Db 517 GTCGGTGGGACGTTATCAAC 537

RESULT 14  
 AX785078  
 LOCUS AX785078 1572 bp DNA linear PAT 17-JUL-2003  
 DEFINITION Sequence 27 from Patent WO03050287.  
 ACCESSION AX785078  
 VERSION AX785078.1 GI:32952909  
 KEYWORDS Arabidopsis thaliana (thale cress)  
 SOURCE Arabidopsis thaliana  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1  
 AUTHORS Schmulling, T. and Werner, T.  
 TITLE Method for modifying plant morphology, biochemistry and physiology  
 JOURNAL Patent: WO 03050287-A 27 19-JUN-2003;  
 Schmulling, Thomas (DE); Werner, Tomas (DE)

FEATURES  
 source Location/Qualifiers

1..1572  
 /organism="Arabidopsis thaliana"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:3702"

ORIGIN

Alignment Scores:  
 Pred. No.: 707 Length: 1572  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-10-014-101B-39 (1-7) x AX785078 (1-1572)

Qy 1 ValGlyGlyThrLeuSerAsn 7  
 Db 517 GTCGGTGGGACGTTATCAAC 537

RESULT 15  
 AF303979  
 LOCUS AF303979 1572 bp mRNA linear PLN 08-NOV-2000  
 DEFINITION Arabidopsis thaliana cytokinin oxidase (CKX3) mRNA, complete cds.  
 ACCESSION AF303979  
 VERSION AF303979.1 GI:11120509  
 KEYWORDS Arabidopsis thaliana (thale cress)  
 SOURCE Arabidopsis thaliana  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 1572)  
 AUTHORS Bilyeu, K.D., Laskey, J.G., Riekhof, W.R., VanVickle, S. and Morris, R.O.

TITLE A family of cytokinin oxidases from Arabidopsis thaliana  
 JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1572)

AUTHORS Bilyeu, K.D., Laskey, J.G., Riekhof, W.R., VanVickle, S. and Morris, R.O.

TITLE Direct Submission

JOURNAL Submitted (11-SEP-2000) University of Missouri, 210 Waters Hall,  
 Columbia, MO 65211, USA

FEATURES Location/Qualifiers

source

1..1572  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /chromosome="V"

gene

1..1572  
 /gene="CKX3"  
 /gene="CKX3"  
 /codon\_start=1  
 /product="cytokinin oxidase"  
 /protein\_id="AAG30906.1"  
 /db\_xref="GI:11120510"

CDS

1..1572  
 /translation="WASYNLSRQVLIATIIIVITLSTPTINTSPQWNILSHNEF  
 AGKLTSSSSVESAAIDFGHVTKIFPSAVLIPSSVEDITLKLKLSFDSQLSPFLAARG  
 HGHSHRQASAKDGVVNMNRDRIKVSRTCLYDVDDAALWIEVLNKLBLEGL  
 TPVSWTDYLYLTGGLTNGGSGQTFRYQITNVLEMDVITGKEIATCKDMNSD  
 LFPAVLGGLGFGIITRAIKLEAVAPKAQWLRFLYIDFSEFRQQRVSKTDGVDF  
 LEGSIMVDHGPDPNWRSTYPPSDHLRIASVMKRRHRYVCLFVVKYDYDSQYTVNEE  
 MEELSDSLNHVRGEMVEXDVTYMDFLNVRTGELNLSKSGQWDVPHWLNLPVKTOI  
 SKPDGQFKGIILRNNTISGVLVYPMRNKNWDRMSAALPEEDVFYAVGFLRSAGFD  
 NWEAFDQENMEILAFCDANMGVIOYLPYHSSQEGWVRHFGFRWNIFVERKYYKPKM  
 ILSQGNIFQKINS"

ORIGIN

Alignment Scores:  
 Pred. No.: 707 Length: 1572  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 8 Gaps: 0

US-10-014-101B-39 (1-7) x AF303979 (1-1572)

Qy 1 ValGlyGlyThrLeuSerAsn 7

Db 517 GTCGGTGGGACGTTATCAAC 537

RESULT 16

AX339731  
 LOCUS AX339731 1575 bp DNA linear PAT 10-JAN-2002  
 DEFINITION Sequence 28 from Patent WO0196580.  
 ACCESSION AX339731  
 VERSION AX339731.1 GI:18135724

KEYWORDS Arabidopsis thaliana (thale cress)

SOURCE Arabidopsis thaliana

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1

AUTHORS Schmulling, T. and Werner, T.

TITLE Method for modifying plant morphology, biochemistry and physiology

JOURNAL Patent: WO 0196580-A 28 20-DEC-2001;

Schmulling, Thomas (DE); Werner, Tomas (DE)

FEATURES Location/Qualifiers

1..1575

/organism="Arabidopsis thaliana"

/mol\_type="unassigned DNA"

/db\_xref="taxon:3702"

ORIGIN

Alignment Scores:

Pred. No.: 708 Length: 1575

Score: 35.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 6 Gaps: 0

US-10-014-101B-39 (1-7) x AX339731 (1-1575)

```

Qy 1 ValGlyGlyThrlleuSerAen 7
Db 535 GTCGGCGGACGTTGTGGAAC 555

RESULT 17
LOCUS AX507394 1575 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 2089 from Patent WO0216655.
ACCESSION AX507394
VERSION AX507394.1 GI:23388631
KEYWORDS Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1
AUTHORS Harper, J.F., Kreps, J., Wang, X. and Zhu, T.
TITLE Stress-regulated genes of plants, transgenic plants containing
same, and methods of use
JOURNAL Patent: WO 0216655-A 2089 28-FEB-2002;
The Scripps Research Institute (US) ; Syngenta Participations AG
(CH)
FEATURES
source Location/Qualifiers
1..1575
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
ORIGIN
Alignment Scores:
Pred. No.: 708 Length: 1575
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-014-101B-39 (1-7) x AX507394 (1-1575)

Qy 1 ValGlyGlyThrlleuSerAen 7
Db 535 GTCGGCGGACGTTGTGGAAC 555

RESULT 18
LOCUS AX651615 1575 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 428 from Patent WO03000898.
ACCESSION AX651615
VERSION AX651615.1 GI:29154433
KEYWORDS Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1
AUTHORS Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
Katagiri, F., Qian, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.
TITLE Plant genes involved in defense against pathogens
JOURNAL Patent: WO 03000898-A 428 03-JAN-2003;
Syngenta Participations AG (CH)
FEATURES
source Location/Qualifiers
1..1575
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
ORIGIN
Alignment Scores:
Pred. No.: 708 Length: 1575
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-014-101B-39 (1-7) x AX651615 (1-1575)

Qy 1 ValGlyGlyThrlleuSerAen 7
Db 535 GTCGGCGGACGTTGTGGAAC 555

RESULT 19
LOCUS AX785079 1575 bp DNA linear PAT 17-JUL-2003
DEFINITION Sequence 28 from Patent WO03050287.
ACCESSION AX785079
VERSION AX785079.1 GI:32952910
KEYWORDS Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1
AUTHORS Schmulling, T. and Werner, T.
TITLE Method for modifying plant morphology, biochemistry and physiology
JOURNAL Patent: WO 03050287-A 28 19-JUN-2003;
Schmulling, Thomas (DE) ; Werner, Tomas (DE)
FEATURES
source Location/Qualifiers
1..1575
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
ORIGIN
Alignment Scores:
Pred. No.: 708 Length: 1575
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-014-101B-39 (1-7) x AX785079 (1-1575)

Qy 1 ValGlyGlyThrlleuSerAen 7
Db 535 GTCGGCGGACGTTGTGGAAC 555

RESULT 20
LOCUS AF303980 1575 bp mRNA linear PLN 08-NOV-2000
DEFINITION Arabidopsis thaliana cytokinin oxidase (CKX4) mRNA, complete cds.
ACCESSION AF303980
VERSION AF303980.1 GI:11120511
KEYWORDS Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 1575)
AUTHORS Bilyeu, K.D., Laskey, J.G., Riekhof, W.R., VanVickle, S. and
Morris, R.O.
TITLE A family of cytokinin oxidases from Arabidopsis thaliana
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1575)
AUTHORS Bilyeu, K.D., Laskey, J.G., Riekhof, W.R., VanVickle, S. and
Morris, R.O.
TITLE Direct Submission
JOURNAL Submitted (11-SEP-2000) University of Missouri, 210 Waters Hall,
Columbia, MO 65211, USA
FEATURES
source Location/Qualifiers
1..1575

```

```

/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="IV"
1..1575
/gene="CKX4"
1..1575
/gene="CKX4"
/codon_start=1
/product="cytokinin oxidase"
/protein_id="AAG30907.1"
/db_xref="GI:11120512"
/translation="MTNTLCLSLITLITPTLISPTLTKSDEGIDVFLPISLNLTLVLT
DPFISASDHDFGNTDENFAGVLCPSSTTEVARLLFFANGFSYNGKSTSPASTPKV
AARQQSHLRQAGAGGVVWMTCLAMAAKPAAVISADGTVDVAAGTMMVDVLKA
AVDRGSPVWTDYLSVGTLSNAGIGGQTFHGQIJSNVHELVDVITGKGMWTC
PKLNPELFGVGLGQGLITRARIADHAPTRKWSRLIYSDFAFKRQDRLISM
TNDGLVFLQGLMMSGVDTSPFLSDQTRVASLVNDHRIIYVLEVAKYDRTLLP
IIDVIDTLKTLGAFGFMVQDVPYDFLNRVNEEDKRLSLGLMEVHPWLNIFV
PGSRIDFHGVINGLLNQTSVGLTFYPTNRKNRMSTWTDVDFYVIGLIQ
SAGSQNQLENLNDKVIQFCNSGIIKIKELMYLHYTRKEDWKHFQKWDVFLRKKI
MFDPKLLSPQDIFN"

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 708 Length: 1575
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0

```

US-10-014-101B-39 (1-7) x AF303980 (1-1575)

```

Qy 1 ValGlyGlyThrLeuSerAsn 7
|||||
Db 535 GTGGGGAGGAGTGTCTCAAC 555

```

## RESULT 21

```

AF303981
LOCUS AF303981 1575 bp mRNA linear PLN 08-NOV-2000
DEFINITION Arabidopsis thaliana cytokinin oxidase (CKX5) mRNA, complete cds.
ACCESSION AF303981
VERSION AF303981.1 GI:11120513
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1575);
Bilyeu,K.D., Laskey,J.G., Riekhof,W.R., VanVickle,S. and
Morris,R.O.
A family of cytokinin oxidases from Arabidopsis thaliana
Unpublished
REFERENCE 2 (bases 1 to 1575)
Bilyeu,K.D., Laskey,J.G., Riekhof,W.R., VanVickle,S. and
Morris,R.O.
Direct Submission
Submitted (11-SEP-2000) University of Missouri, 210 Waters Hall,
Columbia, MO 65211, USA
FEATURES
Location/Qualifiers
1..1575
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="V"
1..1575
/gene="CKX5"
1..1575
/gene="CKX5"
/codon_start=1
/product="cytokinin oxidase"

```

## gene

## CDS

```

/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="V"
1..1575
/gene="CKX4"
1..1575
/gene="CKX4"
/codon_start=1
/product="cytokinin oxidase"

```

```

/protein_id="AAG30908.1"
/db_xref="GI:11120514"
/translation="MIATPEYFLENDAAFAAATAAGKSTDGVSESLNIOEILCGG
AAADTAGRDFGMCNCKVLAIVRPGVPEDIAAGAAKLRSDKLTVAARGNGHSINGQA
MAEGGLVDMSTTAENHFEVGLSGDATAFVDVSGALWEDLVKRCVSEYGLAPRSW
TDYGLTVGGTSLNAGVSGQAFYGPQTSNVTELDVVTGNGDVVTCSEIENSELPFSV
LGLGQFGIITRARILOPAPDMVRIVVTEFDEFTQDAEWLVSQKNESFDFVEG
FYFNGADPVNGWPTVPLPHDHEFPTRLPQSCGVLYLCLEGLHYRDSNSTIDKR
VERLGRLENGSLRPEVDLPYVDFFLRVKRSEEIAGKNGTWTETPHWLNLFVSKRDI
GDFNRTVFKELVKNGVNGLVYPLLRSDRTSVIPEGEI FYI VALLEFVPPCA
KVSVEKVAQNOEIVHWCNKGIDYKLYLPHYKQEEWIRHFGNRSRFDVRKAMFD
PMALLSPQGIKFNRL"

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 708 Length: 1575
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0

```

US-10-014-101B-39 (1-7) x AF303981 (1-1575)

```

Qy 1 ValGlyGlyThrLeuSerAsn 7
|||||
Db 502 GTGGGGAGGAGTGTCTCAAT 522

```

## RESULT 22

```

AF303027
LOCUS AF303027 1587 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 2897 from Patent WO03000898.
ACCESSION AF303027
VERSION AF303027.1 GI:29155841
KEYWORDS Oryza sativa
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzae; Oryza.
1
REFERENCE Chang,H.S., Chen,W., Cooper,B., Glazebrook,J., Goff,S.A., Hou,Y.M.,
AUTHORS Katagiri,F., Quan,S., Tao,Y., Whitham,S., Xie,Z., Zhu,T. and Zou,G.
TITLE Plant genes involved in defense against pathogens
JOURNAL Patent: WO 03000898-A 2897 03-JAN-2003;
SYNGENTA PARTICIPATIONS AG (CH)
FEATURES
Location/Qualifiers
1..1587
/organism="Oryza sativa"
/mol_type="unassigned DNA"
/db_xref="taxon:4530"

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 713 Length: 1587
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

```

US-10-014-101B-39 (1-7) x AX653027 (1-1587)

```

Qy 1 ValGlyGlyThrLeuSerAsn 7
|||||
Db 523 GTGGGTGGCAGCTCTCCAC 543

```

## RESULT 23

```

AX653226
LOCUS AX653226 1590 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 3096 from Patent WO03000898.
ACCESSION AX653226
VERSION AX653226.1 GI:29156040

```

```

KEYWORDS
SOURCE Oryza sativa
ORGANISM Oryza sativa
DEFINITION Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
ACCESSION Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
VERSION Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
AUTHORS Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
TITLE Katagiri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.
JOURNAL Plant genes involved in defense against pathogens
PATENT: WO 03000898-A 3096 03-JAN-2003;
SYNGENTA Participations AG (CH)
FEATURES
source Location/Qualifiers
1..1590
/organism="Oryza sativa"
/mol_type="unassigned DNA"
/db_xref="taxon:4530"
ORIGIN
Alignment Scores:
Pred. No.: 714 Length: 1590
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-014-101B-39 (1-7) x AX653226 (1-1590)
QY 1 ValGlyGlyThrLeuSerAsn 7
Db 520 GTCGGTGGCACCTGTGCAAT 540
RESULT 24
LOCUS AX652883 1593 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 2753 from Patent WO03000898.
ACCESSION AX652883
VERSION AX652883.1 GI:29155697
KEYWORDS
SOURCE Oryza sativa
ORGANISM Oryza sativa
DEFINITION Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
ACCESSION Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
VERSION Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
AUTHORS Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
TITLE Katagiri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.
JOURNAL Plant genes involved in defense against pathogens
PATENT: WO 03000898-A 2753 03-JAN-2003;
SYNGENTA Participations AG (CH)
FEATURES
source Location/Qualifiers
1..1593
/organism="Oryza sativa"
/mol_type="unassigned DNA"
/db_xref="taxon:4530"
ORIGIN
Alignment Scores:
Pred. No.: 715 Length: 1593
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-014-101B-39 (1-7) x AX652883 (1-1593)
QY 1 ValGlyGlyThrLeuSerAsn 7
Db 517 GTGGGGGGCACCTCTTCCAC 537
RESULT 25

AR150952 AR150952 1605 bp DNA linear PAT 08-AUG-2001
LOCUS Sequence 3 from patent US 6229066.
ACCESSION AR150952
VERSION AR150952.1 GI:15115543
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1605)
AUTHORS Morris, R.O.
TITLE Cytokinin oxidase
JOURNAL Patent: US 6229066-A 3 08-MAY-2001;
FEATURES
source Location/Qualifiers
1..1605
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Alignment Scores:
Pred. No.: 719 Length: 1605
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-014-101B-39 (1-7) x AR150952 (1-1605)
QY 1 ValGlyGlyThrLeuSerAsn 7
Db 523 GTCGGGGCACCGCTGTCCAAC 543
RESULT 26
LOCUS AR399211 1605 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 3 from patent US 6617497.
ACCESSION AR399211
VERSION AR399211.1 GI:40137776
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1605)
AUTHORS Morris, R.O.
TITLE Cytokinin oxidase
JOURNAL Patent: US 6617497-A 3 09-SEP-2003;
FEATURES
source Location/Qualifiers
1..1605
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 719 Length: 1605
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-014-101B-39 (1-7) x AR399211 (1-1605)
QY 1 ValGlyGlyThrLeuSerAsn 7
Db 523 GTCGGGGCACCGCTGTCCAAC 543
RESULT 27
LOCUS BT002757 1606 bp mRNA linear PLN 15-JAN-2003
DEFINITION Arabidopsis thaliana CKX5 mRNA sequence.
ACCESSION BT002757
VERSION BT002757.1 GI:27754264

```

**KEYWORDS**  
**SOURCE** Arabidopsis thaliana (thale cress)  
**ORGANISM** Arabidopsis thaliana  
**REFERENCE** Arabidopsis thaliana (thale cress); Streptophyta; Embryophyta; Tracheophyta;  
**AUTHORS** Eukaryota; Viridiplantae; Magnoliophyta; eudicotyledons; core eudicots;  
 1 (bases 1 to 1606) rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
**TITLE** Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M.,  
**JOURNAL** Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C.,  
**REFERENCE** Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M.,  
**AUTHORS** Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W.,  
 Ecker, J.R., and Theologis, A.  
**TITLE** Arabidopsis Open Reading Frame (ORF) Clones  
**JOURNAL** Unpublished  
**REFERENCE** 2 (bases 1 to 1606)  
**AUTHORS** Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M.,  
 Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C.,  
 Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M.,  
 Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W.,  
 Ecker, J.R., and Theologis, A.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (15-JAN-2003) Plant Gene Expression Center, 800 Buchanan  
**COMMENT** Street, Albany, CA 94710, USA  
 Annotation based on July 2002 version of the Arabidopsis genome  
 submitted to Genbank.  
**FEATURES**  
 source  
 1..1606  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /chromosome="4"  
 /clone="C105456"  
 /ecotype="Columbia"  
 /note="This clone is in pUNI 51."  
 1..1606  
 /gene="At4g14745"  
 /gene="At4g14745"  
 /note="This clone does not hit any of the Arabidopsis  
 chromosomes (version July 2002), but it does hit another  
 cDNA entry (Accession Number AF303981 and AY091158)  
 encoding a putative cytokinin oxidase."  
 /evidence="experimental"  
**ORIGIN**  
 Alignment Scores:  
 Pred. No.: 720 Length: 1606  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 8 Gaps: 0  
 US-10-014-101B-39 (1-7) x BT002757 (1-1606)  
**QY** 1 ValGlyGlyThrLeuSerAsn 7  
 |||||  
**Db** 502 GTGGAGGTACGTGTCAAAAT 522  
 RESULT 28  
 LOCUS AX039923 1608 bp DNA linear PAT 18-NOV-2000  
 DEFINITION Sequence 1 from Patent WO0063401.  
 ACCESSION AX039923  
 VERSION AX039923.1 GI:11229952  
 KEYWORDS  
 Zea mays  
 Zea mays  
**ORGANISM** Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.  
**REFERENCE** 1  
**AUTHORS** Habben, J.E., Zinsmeister, C. and Tomes, D.  
 Regulated expression of genes in plant seeds  
 Patent: WO 0063401-A 1 26-OCT-2000;  
 PIONEER HI-BRED INTERNATIONAL, INC. (US)  
**FEATURES**  
 source  
 1..1608  
 /organism="Zea mays"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:4577"  
 1..1608  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="C11229953"  
 /db\_xref="GI:11229953"  
 /translation="NAVYVYLLAGLIACSHALAGTALGEDGRGWPAPLAALALD  
 GKLPDTSNATAASTDFGNTISALPAAVLYPSSTGDLVALISAANSTGWPYTIAPRG  
 RGHSLMGAPAPGCVVNNASIGDAAAPRVNVSADGRVVDAGGEQVWIDVLRASL  
 ARGVAPRSMTDLYLTVTGGLTNSAGISQAFRHHGQFQISNVLEMDVITGHGEMVTCCKQ  
 LNADIFDVLGLGQFGVITRAVEAPAPARWRLVYTDFAAFSAQDRLTAPRP  
 GGGASFGPMSYVEGSFVNQSLATDLANTGFTDADVARVALAGERNATVYSIEA  
 TLVYDNATAAAADQELASVLGTSYVEGFAFQDVSVTAFLDRVHGEEVALNKLGLW  
 RVPHPLNMFPVRSRIADFDRGVEKGIQGTDIYGLVPLVYPLNKMDDGMSAATPSE  
 DVFYAVSLFSSVAPNDLARLQQRNRRLRFCDLACIOYKTYLAKHTDRSDWVRHFGA  
 AEWNRFEVKNKIDPRLLSPGQDIFN"  
**ORIGIN**  
 Alignment Scores:  
 Pred. No.: 721 Length: 1608  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0  
 US-10-014-101B-39 (1-7) x AX039923 (1-1608)  
**QY** 1 ValGlyGlyThrLeuSerAsn 7  
 |||||  
**Db** 529 GTGGCGGCGACGCTGTCCAAC 549  
 RESULT 29  
 LOCUS AX339732 1611 bp DNA linear PAT 10-JAN-2002  
 DEFINITION Sequence 29 from Patent WO0196580.  
 ACCESSION AX339732  
 VERSION AX339732.1 GI:18135725  
 KEYWORDS  
 Arabidopsis thaliana (thale cress)  
**ORGANISM** Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
**REFERENCE** 1  
**AUTHORS** Schmullling, T. and Werner, T.  
**TITLE** Method for modifying plant morphology, biochemistry and physiology  
**JOURNAL** Patent: WO 0196580-A 29 20-DEC-2001;  
 Schmullling, Thomas (DE); Werner, Tomas (DE)  
**FEATURES**  
 source  
 1..1611  
 /organism="Arabidopsis thaliana"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:3702"  
**ORIGIN**  
 Alignment Scores:  
 Pred. No.: 722 Length: 1611  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0  
 US-10-014-101B-39 (1-7) x AX339732 (1-1611)

QY 1 ValGlyGlyThrLeuSerAsn 7  
Db 499 GTTGGAGGTACTCTCCAAT 519

RESULT 30  
LOCUS AX785080 1611 bp DNA linear PAT 17-JUL-2003  
DEFINITION Sequence 29 from Patent WO03050287.  
ACCESSION AX785080  
VERSION AX785080.1 GI:32952911  
KEYWORDS Arabidopsis thaliana (thale cress)  
SOURCE Arabidopsis thaliana  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1  
AUTHORS Schmulling, T. and Werner, T.  
TITLE Method for modifying plant morphology, biochemistry and physiology  
JOURNAL Patent: WO 03050287-A 29 19-JUN-2003;  
Schmulling, Thomas (DE); Werner, Tomas (DE)  
FEATURES  
source  
1. .1611  
/organism="Arabidopsis thaliana"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3702"

ORIGIN  
Alignment Scores:  
Pred. No.: 722 Length: 1611  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-014-101B-39 (1-7) x AX785080 (1-1611)

QY 1 ValGlyGlyThrLeuSerAsn 7  
Db 499 GTTGGAGGTACTCTCCAAT 519

RESULT 31  
LOCUS AX339737 1620 bp DNA linear PAT 10-JAN-2002  
DEFINITION Sequence 34 from Patent WO0196580.  
ACCESSION AX339737  
VERSION AX339737.1 GI:18135729  
KEYWORDS Arabidopsis thaliana (thale cress)  
SOURCE Arabidopsis thaliana  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1  
AUTHORS Schmulling, T. and Werner, T.  
TITLE Method for modifying plant morphology, biochemistry and physiology  
JOURNAL Patent: WO 0196580-A 34 20-DEC-2001;  
Schmulling, Thomas (DE); Werner, Tomas (DE)  
FEATURES  
source  
1. .1620  
/organism="Arabidopsis thaliana"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3702"

ORIGIN  
Alignment Scores:  
Pred. No.: 725 Length: 1620  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-014-101B-39 (1-7) x AX339737 (1-1620)

QY 1 ValGlyGlyThrLeuSerAsn 7  
Db 508 GTTGGAGGTACTCTCCAAT 528

RESULT 32  
LOCUS AX785085 1620 bp DNA linear PAT 17-JUL-2003  
DEFINITION Sequence 34 from Patent WO03050287.  
ACCESSION AX785085  
VERSION AX785085.1 GI:32952915  
KEYWORDS Arabidopsis thaliana (thale cress)  
SOURCE Arabidopsis thaliana  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1  
AUTHORS Schmulling, T. and Werner, T.  
TITLE Method for modifying plant morphology, biochemistry and physiology  
JOURNAL Patent: WO 03050287-A 34 19-JUN-2003;  
Schmulling, Thomas (DE); Werner, Tomas (DE)  
FEATURES  
source  
1. .1620  
/organism="Arabidopsis thaliana"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3702"

ORIGIN  
Alignment Scores:  
Pred. No.: 725 Length: 1620  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-014-101B-39 (1-7) x AX785085 (1-1620)

QY 1 ValGlyGlyThrLeuSerAsn 7  
Db 508 GTTGGAGGTACTCTCCAAT 528

RESULT 33  
LOCUS AF303982 1623 bp mRNA linear PLN 08-NOV-2000  
DEFINITION Arabidopsis thaliana cytokinin oxidase (CKX6) mRNA, complete cds.  
ACCESSION AF303982  
VERSION AF303982.1 GI:11120515  
KEYWORDS Arabidopsis thaliana (thale cress)  
SOURCE Arabidopsis thaliana  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 1623)  
AUTHORS Bilyeu, K.D., Laskey, J.G., Riekhof, W.R., VanVickle, S. and Morris, R.O.  
TITLE A family of cytokinin oxidases from Arabidopsis thaliana  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1623)  
AUTHORS Bilyeu, K.D., Laskey, J.G., Riekhof, W.R., VanVickle, S. and Morris, R.O.  
TITLE Direct Submission  
JOURNAL Submitted (11-SEP-2000) University of Missouri, 210 Waters Hall,  
Columbia, MO 65211, USA  
FEATURES  
source  
1. .1623  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"

```

gene      /chromosome="1"
1..1623
/gene="CKX6"
CDS       1..1623
/gene="CKX6"
/codon_start=1
/product="cytokinin oxidase"
/protein_id="AAG30909.1"
/db_xref="GI:11120516"
/translation="MREMTSSFLITFAICKLIIAVGLNVGPSELLRICAIIDVDGHP
TVHPSDLASVSSDPGMLKSPPEPLAVLHPSAEDVARLVRTAIGSATFPVSARGHGH
SINGQAAAGRNQVVENMNGVITGPFLVPDPMYVDVVGGLWLVVLLKTLHGLAP
KSTWIDLYLVGGTSLNAGISGQALHHPQISNVLEDDVVTGKGEVMRCEEENTRLF
HGVLGGGQGIITRARSISLEPAPQVRWIRVLYSFKVFTQDEYLIISHGOLKFDY
VEGFVIIVDEGLVNNWRSFSPNPKVSISSVNGSVLYCLEITKNYHSDSIVDQE
VEILMKLNIPVSTVFTDLYQVDFLDVHKAELKLRSLNVEVPHPLNLFVFKSRI
SDPKGVFKGLGNKTSGLPIIYPMKQKWDERSAVTPDEEVFYIIVALLRSALTQGE
ETQKLEVLKQNRILIEFCQAKINVKYILPHHATQEEWVAHFQDKWDRFRSLKAEFD
FRHLATGQRIFQNPSSLSPFPSSSSSAASW"

ORIGIN
Alignment Scores:
Pred. No.: 726 Length: 1623
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0

US-10-014-101B-39 (1-7) x AF303982 (1-1623)

Qy 1 ValGlyGlyThrLeuSerAsn 7
Db 511 GTTGGAGGTACACTCTCCAAT 531

RESULT 34
LOCUS BT000179 1655 bp mRNA linear PLN 19-SEP-2002
DEFINITION Arabidopsis thaliana cytokinin oxidase-like protein (At4g29740)
mRNA, complete cds.
ACCESSION BT000179
VERSION BT000179.1 GI:23197941
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1655)
Nguyen,M., Karlin-Neumann,G., Southwick,A., Tripp,M., Miranda,M.,
Palm,C.J., Bowser,L., Jones,T., Barth,J., Carninci,P., Chen,H.,
Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,
Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,
Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K.,
Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submission
Submitted (19-SEP-2002) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
e-mail for correspondence: arab@sequence.stanford.edu

RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAPL cDNAs (RAPL cDNA : 'RIKEN
Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PDEC (SSP) Consortium members constructed and
sequenced the PENTR (ORF) clones using the RAPL cDNAs: Nguyen,M.,
Southwick,A., Tripp,M., Palm,C.J., Jones,T., Wu,T., Chen,H.,
Cheuk,R., Chan,M.M., Chang,C.H., Dale,J.M., Deng,J.M., Hsuan,V.W.,
Lee,J.M., Kim,C.J., Quach,H.L., Shinn,P., Tang,C.C., Toroumi,M.,
Wallender,E.K., Wong,C., Wu,H.C., Yamada,K., Yu,G., Yuan,S.,

FEATURES
source
1..1655
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="4"
/clone="U16213"
/ecotype="Columbia"
/Note="this clone is in PENTR/SD-dTopo This is a cloned
PCR product using RIKEN clone RAFL07-11-K12 (AY054460) as
a template"
gene 1..1655
/gene="At4g29740"
/Note="synonym: T16L4.250"
CDS 1..1575
/gene="At4g29740"
/codon_start=1
/product="cytokinin oxidase-like protein"
/protein_id="AAN15498.1"
/db_xref="GI:23197942"
/translation="MTNTLCISLITLTLFISLTPTLIKSDGIDVFLPISNLTVLTL
DPFISAAASHDPMITDENPGAVLCPSTTEVARLLRFANGSGFYKNGKSTSPASTFKV
AARGQHSIRGQASAGGVVNMTCIAWAAPAAVVISADGTIADVAAGTMMVDVKA
AVDGVSPVTWTDLYLSVGGTGNAGIGGQTFRRGPOISNVHEDLVITGGEWMTCS
KLNLPFLVGLGQFGIITRARIADHAPTRVKSRIILYSDFSAFKRYQDRTLLSM
TNDGLVLEGLQAMNSGFVDTSPFSDQTRVASLVNDHRIIYVLEVAKYDRTLLP
IIDVIDTLRTLGAPGFVQDVPYDFPLNRVNERDKLSLGLWEVPHPLNLFV
PGRSQDFHDGVLNGLINOTSTSGVTILFPTNKNKNRMSTMTDDEVPYVIGLLQ
SAGSQWQLELENLNDKVIQFCENSGIKIKLYLMHYTRKEDVWVHFQDWDFFLRKKI
MDFPKRLSLSPGQDIFN"

ORIGIN
Alignment Scores:
Pred. No.: 738 Length: 1655
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0

US-10-014-101B-39 (1-7) x BT000179 (1-1655)

Qy 1 ValGlyGlyThrLeuSerAsn 7
Db 535 GTCGCGGCGAGCTTGTGCGAAC 555

RESULT 35
LOCUS AX553630 1677 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 3500 from Patent WO03000898.
ACCESSION AX553630
VERSION AX553630.1 GI:29156444
KEYWORDS Oryza sativa
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1
Chang,H.S., Chen,W., Cooper,B., Glazebrook,J., Goff,S.A., Hou,Y.M.,
Katagiri,F., Qian,S., Tao,Y., Whitham,S., Xie,Z., Zhu,T. and Zou,G.
Plant genes involved in defense against pathogens
Patent: WO 03000898-A 3500 03-JAN-2003;
Syngenta Participations AG (CH)
Location/Qualifiers
1..1677
/organism="Oryza sativa"
/mol_type="unassigned DNA"

```

/db\_xref="taxon:4530"

## ORIGIN

Alignment Scores:  
 Pred. No.: 746 Length: 1677  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-10-014-101B-39 (1-7) x AX653630 (1-1677)

Qy 1 ValGlyGlyThrLeuSerAsn 7

Db 532 GTCGGCGGACGCTGTCCAAC 552

## RESULT 36

BT004107 1687 bp mRNA linear PLN 14-FEB-2003  
 LOCUS Arabidopsis thaliana clone RAFL15-29-H04 (R20989) putative  
 DEFINITION cytochrome oxidase (At2g19500) mRNA, complete cds.

ACCESSION BT004107

VERSION BT004107.1 GI:28393415

KEYWORDS FLI CDNA.

SOURCE Arabidopsis thaliana (thale cress)

## ORGANISM

Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

1 (bases 1 to 1687)

Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Huan, V.W., Lee, J.M.,  
 Onodera, C.S., Quach, H.L., Tang, C., Toriumi, M., Wong, C., Wu, H.C.,  
 Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y.,  
 Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M.,  
 Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P.,  
 Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W.,  
 Ecker, J.R. and Theologis, A.  
 Arabidopsis Full Length cDNA Clones

Unpublished

2 (bases 1 to 1687)

Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Huan, V.W., Lee, J.M.,  
 Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C.,  
 Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y.,  
 Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M.,  
 Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P.,  
 Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W.,  
 Ecker, J.R. and Theologis, A.

Direct Submission

Submitted (14-FEB-2003) Plant Gene Expression Center, 800 Buchanan  
 Street, Albany, CA 94710, USA

RIKEN Genomic Sciences Center (GSC) members carried out the

collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN

Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J.,

Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,

Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the  
 sequencing and annotation of the RAFL cDNAs: Yamada, K., Chan, M.M.,  
 Chang, C.H., Dale, J.M., Huan, V.W., Lee, J.M., Onodera, C.S.,  
 Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yu, G.,  
 Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M.,  
 Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W.,  
 Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to  
 this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC)  
 contributed equally to this work as PIs.

Annotation based on July 2002 version of the Arabidopsis genome  
 submitted to Genbank.

Location/Qualifiers

1. .1687

## FEATURES

-source

/organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /chromosome="2"  
 /clone="RAFL15-29-H04 (R20989)"  
 /ecotype="Columbia"  
 /note="this clone is in a modified pBluescript vector2  
 (lambda PS) as a BamHI/XhoI insert."  
 1. .1687  
 /gene="At2g19500"  
 1. .54  
 /gene="At2g19500"  
 55. .1560  
 /gene="At2g19500"  
 /codon\_start=1  
 /evidence=experimental  
 /product="putative cytochrome oxidase"  
 /protein\_id="AA042130.1"  
 /db\_xref="GI:28393415"  
 /translation="MANRLMTLITVLMITKSSNGIKIDLPKSLNLTSLTSDPSIIISA  
 ASHDFGNITVTPGGVICPSSTADISRLQYAAAGKSTFQVARGQGHSLNGQASVSG  
 GIVNMTCITDVVSKKAYADAAAGTLDVLTAKTAEKGVSPVTLHITVGTGL  
 SNGGIGGVFRNGPLVNLDELDTITKQERLISANDINPELFGVGLGIGQFGIITR  
 ARVLDHAPKRAKWFRLYDFTFTKQERLISANDINPELFGVGLGIGQFGIITR  
 FFPDSQKADLVKQHGIIYVLEAKYDDNLPITISKVIDTLTKTSLVLPQFISMH  
 DVAYDFLNRVHVENKLSGLWELPHEPWLNLVYVPSKRLDFHNGVWVKLLKQKSA  
 SGLALYPTNRKNWNRMSAMIPEDVDVITIGLQSATPKDLPVPSVNEKIIIRFC  
 KDSGKIKQYLMHTTSKEDIHFSGKNDPDKRDLFPDKLLSPQDIF"  
 1561. .1687  
 /gene="At2g19500"

## ORIGIN

Alignment Scores:  
 Pred. No.: 750 Length: 1687  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 8 Gaps: 0

US-10-014-101B-39 (1-7) x BT004107 (1-1687)

Qy 1 ValGlyGlyThrLeuSerAsn 7

Db 520 GTCGGAGGACGTTGCGAAT 540

## RESULT 37

CQ812635 1728 bp DNA linear PAT 24-MAY-2004  
 LOCUS Sequence 1 from Patent WO2004038027.  
 DEFINITION CQ812635

ACCESSION CQ812635

VERSION CQ812635.1 GI:47602085

KEYWORDS Arabidopsis thaliana (thale cress)

SOURCE Arabidopsis thaliana

Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

1

van Camp, W.

Bioremediation

Patent: WO 2004038027-A 1 06-MAY-2004;

CropDesigN N.V. (BE)

Location/Qualifiers

1. .1728

/organism="Arabidopsis thaliana"

/mol\_type="unassigned DNA"

/db\_xref="taxon:3702"

## ORIGIN

Alignment Scores:

Pred. No.: 766 Length: 1728

Score: 35.00 Matches: 7

```

Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 6                   Indels: 0
DB:                               Gaps: 0

US-10-014-101B-39 (1-7) x CQ812635 (1-1728)

QY 1 ValGlyGlyThrLeuSerAsn 7
Db 574 GTTGAGGTACACTATCTAAT 594

RESULT 38
AX339728 1728 bp DNA linear PAT 10-JAN-2002
LOCUS
DEFINITION Sequence 25 from Patent WO0196580.
ACCESSION AX339728
VERSION AX339728.1 GI:18135721
KEYWORDS Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1
AUTHORS Schmullling,T. and Werner,T.
TITLE Method for modifying plant morphology, biochemistry and physiology
JOURNAL Patent: WO 0196580-A 25 20-DEC-2001;
Schmullling, Thomas (DE) ; Werner, Tomas (DE)
FEATURES
source
1..1728
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Alignment Scores:
Pred. No.: 766 Length: 1728
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6 Indels: 0
DB: Gaps: 0

US-10-014-101B-39 (1-7) x AX339728 (1-1728)

QY 1 ValGlyGlyThrLeuSerAsn 7
Db 574 GTTGAGGTACACTATCTAAT 594

RESULT 39
AX785076 1728 bp DNA linear PAT 17-JUL-2003
LOCUS
DEFINITION Sequence 25 from Patent WO03050287.
ACCESSION AX785076
VERSION AX785076.1 GI:32952907
KEYWORDS Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1
AUTHORS Schmullling,T. and Werner,T.
TITLE Method for modifying plant morphology, biochemistry and physiology
JOURNAL Patent: WO 03050287-A 25 19-JUN-2003;
Schmullling, Thomas (DE) ; Werner, Tomas (DE)
FEATURES
source
1..1728
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN

```

```

Alignment Scores:
Pred. No.: 766 Length: 1728
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6 Indels: 0
DB: Gaps: 0

US-10-014-101B-39 (1-7) x AX785076 (1-1728)

QY 1 ValGlyGlyThrLeuSerAsn 7
Db 574 GTTGAGGTACACTATCTAAT 594

RESULT 40
ZMY18377 1776 bp mRNA linear PLN 10-MAY-1999
LOCUS
DEFINITION Zea mays mRNA for cytokinin oxidase.
ACCESSION Y18377
VERSION Y18377.1 GI:3882017
KEYWORDS cytokinin oxidase.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1
AUTHORS Houben-Herin,N., Pethe,C., d'Alayer,J. and Laloue,M.
TITLE Cytokinin oxidase from Zea mays: purification, cDNA cloning and
expression in moss protoplasts
JOURNAL Plant J. 17 (6), 615-626 (1999)
MEDLINE 99246676
PUBMED 10230061
REFERENCE 2 (bases 1 to 1776)
AUTHORS Laloue,M.
TITLE Direct Submission
JOURNAL Submitted (12-NOV-1998) M. Laloue, INRA Laboratoire de Biologie
Cellulaire, Route de Saint-Cyr, F-78026 Versailles Cedex, FRANCE
COMMENT Conflicting sequence: AF044603.
FEATURES
source
1..1776
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="nobilis"
/db_xref="taxon:4577"
/dev_stage="kernel"
1..1776
/gene="cko"
38..1642
/gene="cko"
/codon_start=1
/product="cytokinin oxidase"
/protein_id="CAA77151.1"
/db_xref="GI:3882018"
/db_xref="GOA:Q9TON8"
/db_xref="UniProt/Swiss-Prot:O9TON8"
/translation="MAVYVYLLAGLIACSHALAGTTPALGDGDRGWPASIALALD
GKLTDSNATAASTDFGNITSAIPAAVLFPSTADIVALLSAANSTPGWYTTAFRG
RGHSLMQAPAGPGVYVNMASLGDAAPPRINVSAGRYVDVAGGEQWIDVLRASLR
GVAPRWTDYLYLTGVTLSNAGISGQAFRHGPQISNVLEMDVITGHGEMVTCSQLN
ADLFDAVLGGVGQFVITRAIAYEPAPARWRVLYVTDFAAFSADQERLTVPKPG
GGASFGPMSSVEGVSFVNQSLATLANTGFTDADVAIVFALAGERNATTVYSTEATL
NYDNATAAADVDDQLSLTSLYVEGFAFQFDVAFAFLDRVHGEVALNKLGLWR
VPHPLANFVPRSIADFDGRGVKGIQGTDIIVGLPIVPLNKMWDGMSAATPSED
VFVAVSLFSSVAPNDLARLQEQNRRLRCDLGIQYKTYLARHTDSDWRHFGAA
KWRVFEMKNDKIDPRLLSPGQDIFN"

ORIGIN
Alignment Scores:
Pred. No.: 783 Length: 1776
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

```

Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0  
US-10-014-101B-39 (1-7) x ZMY18377 (1-1776)  
QY 1 ValGlyGlyThrLeuSerAsn 7  
| | | | | | | | | | | | | | | | | | | | |  
Db 560 GTGGGGGACGGTGTCCAC 580

Search completed: February 18, 2005, 05:21:31  
Job time : 963.84 secs

**This Page Blank (uspto)**

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	35	100.0	342	8	ABZ56423	Abz56423 Aspergill
2	35	100.0	442	3	AC02606	AC02606 Human sc
3	35	100.0	566	10	ADH61273	Adh61273 Soybean c
4	35	100.0	579	13	ACN59295	Acn59295 Cotton gy
5	35	100.0	641	13	ACN54124	Acn54124 Cotton an

XX WPI; 2003-046817/04.  
 XX Detection of expression of specific *Aspergillus* genes for monitoring the  
 PT fermentation and growth conditions of the fungus, using DNA probes.  
 XX  
 XX Claim 1; SEQ ID NO 5536; 48pp + Sequence Listing; Japanese.  
 XX  
 CC The invention relates to a polynucleotide having any of 6006 specific  
 CC sequences (ABZ50888-ABZ56893), which are expressed by a fungus under  
 CC specific culture conditions including one or more of eutrophic,  
 CC oligotrophic, solid, early germination, alkaline, high temperature, low  
 CC temperature or maltose culture or polynucleotides stringently hybridising  
 CC to these sequences. The polynucleotides are useful for monitoring the  
 CC progress of fermentation and the growth conditions of a fungus,  
 CC especially of *Aspergillus oryzae* which is widely used in industrial  
 CC fermentation. Also monitoring for fungal contamination. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pat\_sequences  
 XX  
 SQ Sequence 342 BP; 97 A; 79 C; 88 G; 78 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 99.5 Length: 342  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 8 Gaps: 0  
 US-10-014-101B-39 (1-7) x ABZ56423 (1-342)  
 Qy 1 ValGlyGlyThrLeuSerAsn 7  
 Db 85 GTTGGAGGACACTGAGCAAT 105  
 RESULT 2  
 AAC02606/c  
 ID AAC02606 standard; cDNA; 442 BP.  
 XX  
 AC AAC02606;  
 XX  
 XX 06-OCT-2000 (first entry)  
 DT Human secreted protein 5' EST, SEQ ID NO: 2604.  
 DE  
 DE Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 KW gene therapy; chromosome mapping; ss.  
 KW Homo sapiens.  
 OS  
 XX  
 XX EP1033401-A2.  
 PN  
 XX 06-SEP-2000.  
 XX  
 XX 21-FEB-2000; 2000BP-00200610.  
 PF  
 XX 26-FEB-1999; 99US-0122487P.  
 PR  
 XX (GEST ) GENSET.  
 PA  
 XX Dumas Milne Edwards J, Duclert A, Giordano J;  
 PI WPI; 2000-500381/45.  
 XX P-PSDB; AAG02600.  
 DR  
 XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures.  
 PT  
 XX Claim 1; SEQ ID NO 2604; 71pp + Sequence Listing; English.  
 PS  
 XX

CC The present sequence is one of a large number of 5' ESTs derived from  
 CC mRNAs encoding secreted proteins. An ORF has been identified within the  
 CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs  
 CC derived from 30 different tissues. EST sequences usually correspond  
 CC mainly to the 3' untranslated region (UTR) of the mRNA because they are  
 CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not  
 CC well suited for isolating cDNA sequences derived from the 5' ends of  
 CC mRNAs and even in those cases where longer cDNA sequences have been  
 CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from  
 CC mRNAs with intact 5' ends and can therefore be used to obtain full length  
 CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,  
 CC gene therapy and chromosome mapping procedures. They are used to obtain  
 CC upstream regulatory sequences and to design expression and secretion  
 CC vectors  
 XX  
 SQ Sequence 442 BP; 154 A; 76 C; 116 G; 95 T; 0 U; 1 Other;  
 Alignment Scores:  
 Pred. No.: 132 Length: 442  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 3 Gaps: 0  
 US-10-014-101B-39 (1-7) x AAC02606 (1-442)  
 Qy 1 ValGlyGlyThrLeuSerAsn 7  
 Db 67 GTGGCGGGAACGCTTCTAAT 47  
 RESULT 3  
 ADH61273  
 ID ADH61273 standard; DNA; 566 BP.  
 XX  
 AC ADH61273;  
 XX  
 DT 25-MAR-2004 (first entry)  
 XX  
 DE Soybean cytokinin oxidase (CKX1) DNA #2.  
 DE  
 KW Male-sterility; cytokinin oxidase; transgenic plant; seed production;  
 KW CKX1; soybean; ds; gene.  
 KW  
 OS Glycine max.  
 XX  
 XX US2003163847-A1.  
 PN  
 XX 28-AUG-2003.  
 PD  
 XX 20-DEC-2002; 2002US-00326184.  
 PF  
 XX 20-DEC-2001; 2001US-0343129P.  
 PR  
 XX (PHAA ) PHARMACIA CORP.  
 PA  
 XX Huang S, Crossland LD, Cheikh N, Morris RO;  
 PI WPI; 2003-897983/82.  
 XX GENEANK; BU084470.  
 DR  
 XX Producing plants characterized by reversible male-sterility, useful for  
 PT maintaining male sterility in plants, by transforming a plant cell with a  
 PT nucleic acid construct containing a polynucleotide encoding a cytokinin  
 PT oxidase.  
 XX  
 PS Disclosure; SEQ ID NO 14; 33pp; English.  
 XX  
 XX The invention relates to a method for producing a plant characterised by  
 CC reversible male-sterility which involves transforming a plant cell with a  
 CC nucleic acid construct containing a polynucleotide encoding a cytokinin  
 CC oxidase. The method is useful for producing reversible male-sterility in  
 CC transgenic plants, or for maintaining male sterility in plants. The

CC method reduces the expense of seed production for existing hybrid plants  
 CC such as corn, but also makes it possible to produce hybrid varieties of  
 CC traditionally non-hybrid crops. The method is also useful for introducing  
 CC economically valuable traits from plants having undesirable production  
 CC characteristics into plants having desirable characteristics. The present  
 CC sequence is soybean cytokinin oxidase (CKX1) DNA. This sequence is used  
 CC to illustrate the method of the invention.

SQ Sequence 566 BP; 162 A; 115 C; 132 G; 157 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 174 Length: 566  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 10 Gaps: 0

US-10-014-101B-39 (1-7) x ADH61273 (1-566)

QY 1 ValGlyGlyThrIeuSerAen 7  
 |||||  
 DB 133 GTGGAGGACACTTCCAAAT 153

#### RESULT 4

ACN59295  
 ID ACN59295 standard; cDNA; 579 BP.

AC ACN59295;

DT 02-DEC-2004 (first entry)

XX Cotton gynoeecium tissue EST Clone ID: LIB3829-029-Q6-K6-G4, SEQ:14076.  
 XX Cotton; plant; EST; expressed sequence tag; transgenic plant; gynoeecium;  
 KW variety Nucotton33B; library LIB3829; molecular tag; molecular marker;  
 KW genetic mapping; molecular mapping; seed germination; plant growth;  
 KW plant quality; plant yield; plant breeding; tissue printing; ss.  
 XX Gossypium hirsutum.

OS Gossypium hirsutum.

PN US2004123340-A1.

XX 24-JUN-2004.

XX 12-DEC-2001; 2001US-00021323.

XX 14-DEC-2000; 2000US-0255619P.

XX (DEIK/) DEIKMAN J.

PA (FENG/) FENG P C C.

PA (FINC/) FINCHER K L.

PA (ZIEG/) ZIEGLER T E.

XX Deikman J, Feng PCC, Fincher KL, Ziegler TE;

XX WPI; 2004-479808/45.

XX New isolated nucleic acid molecule that encodes a plant protein or its  
 PT fragment, useful for isolating a variety of agronomically significant  
 PT genes associated with plant growth, quality or yield, and as molecular  
 PT tags to map genes.

XX Claim 1; SEQ ID NO 14076; 34pp; English.

XX The invention relates to 17880 cotton expressed sequence tags (ESTs;  
 CC ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated  
 CC from primed or non-primed seeds from variety DP508, mature seeds from  
 CC variety Coker 312 Boswell 96 Field, and androecium tissue, gynoeecium  
 CC tissue, developing fibres, carpel walls and septa from variety  
 CC Nucotton33B. The invention also relates to substantially purified  
 CC proteins or their fragments encoded by nucleic acid molecules of the  
 CC invention, and to transformed plants having a nucleic acid construct

CC comprising a nucleic acid of the invention. The cotton ESTs are useful as  
 CC molecular tags to isolate genetic regions, to isolate genes, to map  
 CC genes, to determine gene function and to determine whether genes are  
 CC members of a particular gene family. The nucleic acid molecules may be  
 CC used for isolating a variety of agronomically significant genes  
 CC associated with plant growth, quality, yield, and could also serve as  
 CC links in metabolic and catabolic pathways. The nucleic acid molecules are  
 CC also useful for identifying genes important in initiating and maintaining  
 CC seed germination or that may be used to mitigate stresses encountered  
 CC during seed germination. The ESTs additionally enable the acquisition of  
 CC promoters and cis-regulatory elements which will be useful to express  
 CC agronomically significant genes in these tissues and/or other tissues,  
 CC and also permits the acquisition of molecular markers useful in breeding  
 CC schemes, genetic and molecular mapping, and in cloning of agronomically  
 CC significant genes. The nucleic acid molecules are further useful for  
 CC detecting the expression level or pattern of a protein or mRNA and for  
 CC detecting the presence or quantity of a protein by tissue printing. The  
 CC present sequence represents a specifically claimed EST isolated from a  
 CC cotton variety Nucotton33B gynoeecium tissue cDNA library (LIB3829). The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from the US  
 CC patent office at seqdata.uspto.gov/sequence.html?DocID=US20040123340

SQ Sequence 579 BP; 130 A; 136 C; 173 G; 138 T; 0 U; 2 Other;

#### Alignment Scores:

Pred. No.: 178 Length: 579  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 13 Gaps: 0

US-10-014-101B-39 (1-7) x ACN59295 (1-579)

QY 1 ValGlyGlyThrIeuSerAen 7

DB 209 GTTGGTGAACCTGTCTAAC 229

#### RESULT 5

ACN54124

ID ACN54124 standard; cDNA; 641 BP.

XX ACN54124;

XX 02-DEC-2004 (first entry)

XX Cotton androecium tissue EST Clone ID: LIB3828-021-Q1-K6-C8, SEQ:8905.

XX Cotton; plant; EST; expressed sequence tag; transgenic plant; androecium;  
 KW variety Nucotton33B; library LIB3828; molecular tag; molecular marker;  
 KW genetic mapping; molecular mapping; seed germination; plant growth;  
 KW plant quality; plant yield; plant breeding; tissue printing; ss.

OS Gossypium hirsutum.

XX US2004123340-A1.

XX 24-JUN-2004.

XX 12-DEC-2001; 2001US-00021323.

XX 14-DEC-2000; 2000US-0255619P.

XX (DEIK/) DEIKMAN J.

PA (FENG/) FENG P C C.

PA (FINC/) FINCHER K L.

PA (ZIEG/) ZIEGLER T E.

XX Deikman J, Feng PCC, Fincher KL, Ziegler TE;

XX WPI; 2004-479808/45.

XX

PT New isolated nucleic acid molecule that encodes a plant protein or its  
 PT fragment, useful for isolating a variety of agronomically significant  
 PT genes associated with plant growth, quality or yield, and as molecular  
 XX tags to map genes.

XX Claim 1; SEQ ID NO 8905; 34pp; English.

XX The invention relates to 17880 cotton expressed sequence tags (ESTs;  
 CC ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated  
 CC from primed or non-primed seeds from variety DP50B, mature seeds from  
 CC variety Coker 312 Boswell 96 Field, and androecium tissue, gynoecium  
 CC tissue, developing fibres, carpal walls and septa from variety  
 CC Nucleotid3B. The invention also relates to substantially purified  
 CC proteins or their fragments encoded by nucleic acid molecules of the  
 CC invention, and to transformed plants having a nucleic acid construct  
 CC comprising a nucleic acid of the invention. The cotton ESTs are useful as  
 CC molecular tags to isolate genetic regions, to isolate genes, to map  
 CC genes, to determine gene function and to determining whether genes are  
 CC members of a particular gene family. The nucleic acid molecules may be  
 CC used for isolating a variety of agronomically significant genes  
 CC associated with plant growth, quality, yield, and could also serve as  
 CC links in metabolic and catabolic pathways. The nucleic acid molecules are  
 CC also useful for identifying genes important in initiating and maintaining  
 CC seed germination or that may be used to mitigate stresses encountered  
 CC during seed germination. The ESTs additionally enable the acquisition of  
 CC promoters and cis-regulatory elements which will be useful to express  
 CC agronomically significant genes in these tissues and/or other tissues,  
 CC and also permits the acquisition of molecular markers useful in breeding  
 CC schemes, genetic and molecular mapping, and in cloning of agronomically  
 CC significant genes. The nucleic acid molecules are further useful for  
 CC detecting the expression level or pattern of a protein or mRNA and for  
 CC detecting the presence or quantity of a protein by tissue printing. The  
 CC present sequence represents a specifically claimed EST isolated from a  
 CC cotton variety Nucleotid3B androecium tissue cDNA library (LIB3828). The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from the US  
 CC patent office at seqdata.uspto.gov/sequence.html?docid=US20040123340

XX SQ Sequence 641 BP; 147 A; 156 C; 172 G; 166 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 200 Length: 641  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 13 Gaps: 0

US-10-014-101B-39 (1-7) x ACN54124 (1-641)

QY 1 ValGlyGlyThrLeuSerAsn 7  
 Db 541 GTTGGTGAACCTGTGCTAAC 561

RESULT 6  
 ABK28625  
 ID ABK28625 standard; cDNA; 1506 BP.

XX ABK28625;  
 AC ABK28625;  
 XX 09-APR-2002 (first entry)

XX cDNA encoding A. thaliana cytokinin oxidase AtCKX2.

XX Arabidopsis thaliana; plant; cytokinin oxidase; transgenic plant;  
 KW root growth; lateral root; adventitious root; root geotropism; herbicide;  
 KW root meristem; shoot meristem; leaf senescence; parthenocarp; gene; ss.  
 XX Arabidopsis thaliana.

XX WO200196580-A2.

XX 20-DEC-2001.

XX 18-JUN-2001; 2001WO-BP006833.  
 PF  
 XX 16-JUN-2000; 2000EP-00870132.  
 PR 27-DEC-2000; 2000US-0258415P.  
 PR 16-MAR-2001; 2001EP-00870053.

XX (SCHM/) SCHMULLING T.  
 XX (WERN/) WERNER T.

XX Schmulling T, Werner T;

XX WPI; 2002-130736/17.

XX Polynucleotide encoding novel plant protein having cytokinin oxidase  
 PT activity and the protein useful for stimulating root growth, enhancing  
 PT the formation of lateral or adventitious roots, altering root geotropism.  
 XX  
 XX Claim 2; Page 146-147; 154pp; English.

XX The invention relates to an isolated polynucleotide (I) encoding a novel  
 CC plant protein (II) having cytokinin oxidase activity. (I) is useful for  
 CC production of transgenic plants, plant cells or tissues; for production  
 CC of altered plants, plant cell or tissues; and for effecting the  
 CC expression of (II) where (I) is operably linked to one or more control  
 CC sequences. The methods further comprises regenerating a plant from the  
 CC plant cell. (I) and (II) are useful for stimulating root growth;  
 CC enhancing the formation of lateral or adventitious roots; altering root  
 CC geotropism, leading to an increase in yield; and for screening growth  
 CC promoting chemical of herbicides. (I) is useful for increasing the size  
 CC of the root meristem; increasing root size; increasing the size of the  
 CC shoot meristem; delaying leaf senescence and altering leaf senescence;  
 CC increasing leaf thickness; reducing or increasing the vessel size;  
 CC inducing parthenocarp; improving standability of the seedlings;  
 CC increasing branching and for improving lodging resistance. Antibody (III)  
 CC to (II) is useful for identifying and obtaining proteins interacting with  
 CC (II) comprising a screening assay, preferably a two-hybrid screening  
 CC assay. ABK28606-ABK28633 represent Arabidopsis thaliana cytokinin oxidase  
 CC coding sequences and PCR primers of the invention

XX SQ Sequence 1506 BP; 431 A; 328 C; 324 G; 423 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 516 Length: 1506  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-10-014-101B-39 (1-7) x ABK28625 (1-1506)

QY 1 ValGlyGlyThrLeuSerAsn 7  
 Db 466 GTCGGAGGACGTTGTCGAAT 486

RESULT 7  
 ACC85294  
 ID ACC85294 standard; cDNA; 1506 BP.

XX ACC85294;

XX 18-SEP-2003 (first entry)

XX Arabidopsis cytokinin oxidase-like protein 2 cDNA.

XX Maize; root growth; root geotropism; cytokinin oxidase; seed size;  
 KW embryo size; cotyledon size; transgenic plant; herbicide; gene; ss.

XX Arabidopsis thaliana.

XX WO2003050287-A2.

```

PD 19-JUN-2003.
XX
XX 10-DEC-2002; 2002WO-EP013990.
XX
XX 10-DEC-2001; 2001US-00014101.
XX
XX (SCHM/) SCHMULLING T.
XX (WERN/) WERNER T.
XX
XX Schmullling T, Werner T;
XX
XX WPI; 2003-541577/51.
XX
XX Stimulating root growth, enhancing lateral or adventitious root formation
PT or altering root geotropism comprises increasing plant cytokinin oxidase
PT levels or other protein or nucleic acid that reduces active cytokinins in
PT a plant.
XX
XX Claim 3; Page 167-168; 177pp; English.
XX
XX The present invention relates to a method for stimulating root growth or
CC enhancing the formation of lateral or adventitious roots or altering root
CC geotropism, which comprises increasing in a plant or plant part the level
CC of a plant cytokinin oxidase or other protein that reduces the level of
CC active cytokinins in a plant or plant part. Cytokinin oxidase protein and
CC coding sequences from Arabidopsis thaliana are also provided. The method
CC is useful in modifying plant morphological, biochemical and physiological
CC properties, such as in modifying the initiation, stimulation or
CC enhancement of root growth, adventitious root formation, lateral root
CC formation, root geotropism, shoot growth, apical dominance, branching,
CC timing of senescence, timing of flowering, flower formation, seed
CC development and/or seed yield. The present sequence is a coding sequence
XX shown in the invention
XX
SQ Sequence 1506 BP; 431 A; 328 C; 324 G; 423 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 516 Length: 1506
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-014-101B-39 (1-7) x ACC85294 (1-1506)
QY 1 ValGlyGlyThrLeuSerAsn 7
Db 466 GTCGGAGGACGTTGCGAAT 486

RESULT 8
AD006498
ID AD006498 standard; DNA; 1506 BP.
XX
XX AD006498;
AC
XX
XX 29-JUL-2004 (first entry)
XX
XX A thaliana cytokinin oxidase AtCKX2 coding sequence.
XX
XX bioremediation; AtCKX2; cytokinin availability; contaminant;
XX metal deficiency; nutrition; ds; gene; cytokinin oxidase.
XX
XX Arabidopsis thaliana.
XX
XX Key Location/Qualifiers
XX CDS 1..1506
XX /*tag= a
XX /product= "AtCKX2"
XX
XX WO2004038027-A1.
XX
XX 06-MAY-2004.
XX

```

---

```

XX
XX 24-OCT-2003; 2003WO-EP012051.
XX
XX 24-OCT-2002; 2002EP-00079481.
XX
XX (CROP-) CROPDESIGN NV.
XX
XX Van Camp W;
XX
XX WPI; 2004-375913/35.
XX P-PSDB; AD006499.
XX
XX Method for bioremediation, useful for removing contaminants or metals, by
PT decreasing cytokinin availability in plants, and cultivating plant on
PT substrate comprising one or more contaminants.
XX
XX Claim 8; Page 55; 61pp; English.
XX
XX The present invention relates to a method for bioremediation, which
CC involves decreasing cytokinin availability in a plant relative to a
CC corresponding wild type plant, and cultivating the plant on a substrate
CC comprising one or more contaminants, or cultivating a plant having
CC lowered availability of cytokinin relative to corresponding wild type
CC plants, on a substrate, which is to be treated. The method is useful for
CC bioremediation, for concentration of contaminants in a plant, where the
CC plant has a higher concentration of contaminants compared to a
CC corresponding wild type plant. The metal contaminants include aluminum,
CC americium, antimony, arsenic, barium, beryllium, bismuth, cadmium,
CC caesium, cerium, chromium, copper, gallium, germanium, gold, indium,
CC iridium, iron, lead, manganese, mercury, molybdenum, neptunium, osmium,
CC palladium, platinum, plutonium, radium, rhodium, rubidium, selenium,
CC ruthenium, scandium, selenium, silver, strontium, technetium, tellurium,
CC thallium, tin, tungsten, uranium, vanadium or yttrium, preferably
CC cadmium. The plant obtained by the method is useful in bioremediation.
CC Transgenic plants with a lower availability of cytokinin are useful in
CC the manufacture of a medicament for treatment of disorders arising from
CC metal deficiencies, and as a medicament for improving animal or human
CC nutrition. The present sequence is a coding sequence shown in the
XX exemplification of the invention.
XX
SQ Sequence 1506 BP; 431 A; 328 C; 324 G; 423 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 516 Length: 1506
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-10-014-101B-39 (1-7) x AD006498 (1-1506)
QY 1 ValGlyGlyThrLeuSerAsn 7
Db 466 GTCGGAGGACGTTGCGAAT 486

RESULT 9
ABK28629
ID ABK28629 standard; cDNA; 1515 BP.
XX
XX ABK28629;
AC
XX
XX 09-APR-2002 (first entry)
XX
XX cDNA encoding A. thaliana cytokinin oxidase AtCKX6.
XX
XX Arabidopsis thaliana; plant; cytokinin oxidase; transgenic plant;
XX root growth; lateral root; adventitious root; root geotropism; herbicide;
XX root meristem; shoot meristem; leaf senescence; parthenocarp; gene; ss.
XX
XX Arabidopsis thaliana.
XX
XX WO200196580-A2.
XX

```

XX PD 20-DEC-2001.  
 XX XX  
 XX PF 18-JUN-2001; 2001WO-EP006833.  
 XX XX  
 XX PF 16-JUN-2000; 2000EP-00870132.  
 XX PR 27-DEC-2000; 2000US-0258415P.  
 XX PR 16-MAR-2001; 2001EP-00870053.  
 XX XX  
 XX PA (SCHM//) SCHMULLING T.  
 XX PA (WERN//) WERNER T.  
 XX XX  
 XX PI Schmulling T, Werner T;  
 XX XX  
 XX DR WPI; 2002-130736/17.  
 XX XX  
 XX PT Polynucleotide encoding novel plant protein having cytokinin oxidase  
 XX PT activity and the protein useful for stimulating root growth, enhancing  
 XX PT the formation of lateral or adventitious roots, altering root geotropism.  
 XX XX  
 XX PS Example 1; Page 149; 154pp; English.  
 XX XX  
 XX CC The invention relates to an isolated polynucleotide (I) encoding a novel  
 XX CC plant protein (II) having cytokinin oxidase activity. (I) is useful for  
 XX CC production of transgenic plants, plant cells or tissues; for production  
 XX CC of altered plants, plant cell or tissues; and for effecting the  
 XX CC expression of (II) where (I) is operably linked to one or more control  
 XX CC sequences. The methods further comprises regenerating a plant from the  
 XX CC plant cell. (I) and (II) are useful for stimulating root growth;  
 XX CC enhancing the formation of lateral or adventitious root growth;  
 XX CC geotropism, leading to an increase in yield; and for increasing growth  
 XX CC promoting chemical of herbicides. (I) is useful for increasing the size  
 XX CC of the root meristem; increasing root size; increasing the size of the  
 XX CC shoot meristem; delaying leaf senescence and altering leaf senescence;  
 XX CC increasing leaf thickness; reducing or increasing the vessel size;  
 XX CC inducing parthenocarp; improving standability of the seedlings;  
 XX CC increasing branching and for improving lodging resistance. Antibody (III)  
 XX CC to (II) is useful for identifying and obtaining proteins interacting with  
 XX CC (II) comprising a screening assay, preferably a two-hybrid screening  
 XX CC assay. ABK28606-ABK28633 represent Arabidopsis thaliana cytokinin oxidase  
 XX CC coding sequences and PCR primers of the invention  
 XX XX  
 XX SQ Sequence 1515 BP; 456 A; 327 C; 378 G; 354 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 519 Length: 1515  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0  
 US-10-014-101B-39 (1-7) x ABK28629 (1-1515)  
 QY 1 ValGlyGlyThrLeuSerAsn 7  
 DB 493 GTAGGTGGTACTCTGTCCAAAT 513  
 RESULT 10  
 ACC85298  
 ID ACC85298 standard; cDNA; 1515 BP.  
 XX AC  
 XX AC ACC85298;  
 XX XX  
 XX DT 18-SEP-2003 (first entry)  
 XX XX  
 XX DE Arabidopsis cytokinin oxidase-like protein 6 cDNA.  
 XX XX  
 XX KW Maize; root growth; root geotropism; cytokinin oxidase; seed size;  
 XX KW embryo size; cotyledon size; transgenic plant; herbicide; gene; ss.  
 XX OS Arabidopsis thaliana.  
 XX XX

PN WO2003050287-A2.  
 XX XX  
 XX PD 19-JUN-2003.  
 XX XX  
 XX PF 10-DEC-2002; 2002WO-EP013990.  
 XX XX  
 XX PR 10-DEC-2001; 2001US-00014101.  
 XX XX  
 XX PA (SCHM//) SCHMULLING T.  
 XX PA (WERN//) WERNER T.  
 XX XX  
 XX PI Schmulling T, Werner T;  
 XX XX  
 XX DR WPI; 2003-541577/51.  
 XX XX  
 XX PT Stimulating root growth, enhancing lateral or adventitious root formation  
 XX PT or altering root geotropism comprises increasing plant cytokinin oxidase  
 XX PT levels or other protein or nucleic acid that reduces active cytokinins in  
 XX PT a plant.  
 XX XX  
 XX PS Claim 2; Page 171-172; 177pp; English.  
 XX XX  
 XX CC The present invention relates to a method for stimulating root growth or  
 XX CC enhancing the formation of lateral or adventitious roots or altering root  
 XX CC geotropism, which comprises increasing in a plant or plant part the level  
 XX CC of a plant cytokinin oxidase or other protein that reduces the level of  
 XX CC active cytokinins in a plant or plant part. Cytokinin oxidase protein and  
 XX CC coding sequences from Arabidopsis thaliana are also provided. The method  
 XX CC is useful in modifying plant morphological, biochemical and physiological  
 XX CC properties, such as in modifying the initiation, stimulation or  
 XX CC enhancement of root growth, adventitious root formation, lateral root  
 XX CC formation, root geotropism, shoot growth, apical dominance, branching,  
 XX CC timing of senescence, timing of flowering, flower formation, seed  
 XX CC development and/or seed yield. The present sequence is a coding sequence  
 XX CC shown in the invention  
 XX XX  
 XX SQ Sequence 1515 BP; 456 A; 327 C; 378 G; 354 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 519 Length: 1515  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0  
 US-10-014-101B-39 (1-7) x ACC85298 (1-1515)  
 QY 1 ValGlyGlyThrLeuSerAsn 7  
 DB 493 GTAGGTGGTACTCTGTCCAAAT 513  
 RESULT 11  
 AAC43214  
 ID AAC43214 standard; DNA; 1548 BP.  
 XX AC  
 XX AC AAC43214;  
 XX XX  
 XX DT 17-OCT-2000 (first entry)  
 XX XX  
 XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 38443.  
 XX XX  
 XX KW Hybridisation assay; genetic mapping; gene expression control;  
 XX KW protein identification; signal transduction pathway; metabolic pathway;  
 XX KW promoter; termination sequence; ss.  
 XX OS Arabidopsis thaliana.  
 XX XX  
 XX PN EP1033405-A2.  
 XX XX  
 XX PD 06-SEP-2000.  
 XX XX  
 XX PF 25-FEB-2000; 2000EP-00301439.

XX PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0133425P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140931P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 09-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 05-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149168P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 23-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.

```

PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-01571117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158023P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161932P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 532 Length: 1548
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-014-101b-39 (1-7) x AAC43214 (1-1548)
QY 1 ValGlyGlyThrLeuSerAsn 7
DB 466 GTCGGAGGACACTGTCGAAT 486

RESULT 12
ADA71115
ID ADA71115 standard; DNA; 1566 BP.
XX AC ADA71115;
XX DT 20-NOV-2003 (first entry)
XX DE Rice gene, SEQ ID 4438.
XX KW Plant; bacterial infection; fungal infection; viral infection; rice;
XX KW gene; ds.
XX OS Oryza sativa.
XX PN WO2003000898-A1.
XX PD 03-JAN-2003.
XX PF 22-JUN-2001; 2001WO-IB001105.
XX PR 22-JUN-2001; 2001WO-IB001105.
XX XX

(SYGN ) SYNGENTA PARTICIPATIONS AG.
Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,
Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
WPI; 2003-175290/17.

Identifying at least one gene involved in plant resistance or response to
pathogenic infection for conferring resistance or tolerance to a plant to
bacterial, fungal or viral infection by determining or detecting plant
gene expression.

Claim 6; SEQ ID NO 4438; 899pp; English.

The present invention relates to a method (M1) for identifying genes
involved in plant resistance or response to pathogenic infection. M1
comprises identifying a gene whose expression is significantly altered in
the incompatible interaction of plant gene expression relative to
expression of the gene in an uninfected plant, in a mutant plant that
does not express a gene associated with response to pathogenic infection,
or in a corresponding incompatible or compatible interaction. (M1) is
useful for conferring resistance to resistance or tolerance to a plant to
bacterial, fungal or viral infection. The present sequence was used to
illustrate the invention.

SQ Sequence 1566 BP; 422 A; 391 C; 371 G; 382 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 539 Length: 1566
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0

US-10-014-101b-39 (1-7) x ADA71115 (1-1566)
QY 1 ValGlyGlyThrLeuSerAsn 7
DB 499 GTTGGAGGACACTGTCGAAT 519

RESULT 13
ID ABK28626 standard; cDNA; 1572 BP.
XX AC ABK28626;
XX DT 09-APR-2002 (first entry)
XX DE cDNA encoding A. thaliana cytokinin oxidase AtCKX3.
XX KW Arabidopsis thaliana; plant; cytokinin oxidase; transgenic plant;
XX KW root growth; lateral root; adventitious root; root geotropism; herbicide;
XX KW root meristem; shoot meristem; leaf senescence; parthenocarp; gene; ss.
XX OS Arabidopsis thaliana.
XX PN WO200196580-A2.
XX PD 20-DEC-2001.
XX PF 18-JUN-2001; 2001WO-EP006833.
XX PR 16-JUN-2000; 2000EP-00870132.
XX PR 27-DEC-2000; 2000US-0258415P.
XX PR 16-MAR-2001; 2001EP-00870053.
XX XX
(SCHM/) SCHMULLING T.
XX PA (WERN/) WERNER T.
XX PI Schmulling T, Werner T;
XX XX
WPI; 2002-130736/17.

```

XX Polynucleotide encoding novel plant protein having cytokinin oxidase  
PT activity and the protein useful for stimulating root growth, enhancing  
PT the formation of lateral or adventitious roots, altering root geotropism.  
XX  
XX Claim 3; Page 147; 154pp; English.  
XX  
XX The invention relates to an isolated polynucleotide (I) encoding a novel  
CC plant protein (II) having cytokinin oxidase activity. (I) is useful for  
CC production of transgenic plants, plant cells or tissues; for production  
CC of altered plants, plant cell or tissues; and for effecting the  
CC expression of (II) where (I) is operably linked to one or more control  
CC sequences. The methods further comprises regenerating a plant from the  
CC plant cell: (I) and (II) are useful for stimulating root growth;  
CC enhancing the formation of lateral or adventitious roots; altering root  
CC geotropism, leading to an increase in yield; and for screening growth  
CC promoting chemical of herbicides. (I) is useful for increasing the size  
CC of the root meristem; increasing root size; increasing the size of the  
CC shoot meristem; delaying leaf senescence and altering leaf senescence;  
CC increasing leaf thickness; reducing or increasing the vessel size;  
CC inducing parthenocarpy; improving standability of the seedlings;  
CC increasing branching and for improving lodging resistance. Antibody (III)  
CC to (II) is useful for identifying and obtaining proteins interacting with  
CC (II) comprising a screening assay, preferably a two-hybrid screening  
CC assay. ABK28606-ABK28633 represent Arabidopsis thaliana cytokinin oxidase  
CC coding sequences and PCR primers of the invention  
XX  
SQ Sequence 1572 BP; 461 A; 327 C; 348 G; 436 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 541 Length: 1572  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-014-101B-39 (1-7) x ABK28626 (1-1572)

QY 1 ValGlyGlyThrLeuSerAsn 7  
|||||  
DB 517 GTCGGTGGGACGTTATCAAC 537

RESULT 14

ACC85295  
ID ACC85295 standard; cDNA; 1572 BP.

XX ACC85295;

AC 18-SEP-2003 (first entry)

DT Arabidopsis cytokinin oxidase-like protein 3 cDNA.

DE Maize; root growth; root geotropism; cytokinin oxidase; seed size;  
KW embryo size; cotyledon size; transgenic plant; herbicide; gene; ss.

XX Arabidopsis thaliana.

XX WO2003050287-A2.

XX 19-JUN-2003.

XX 10-DEC-2002; 2002WO-EP013990.

XX 10-DEC-2001; 2001US-00014101.

XX (SCHM/) SCHMULLING T.

PA (WERN/) WERNER T.

XX Schmullling T, Werner T;

XX WPI; 2003-541577/51.

PT Stimulating root growth, enhancing lateral or adventitious root formation  
PT or altering root geotropism comprises increasing plant cytokinin oxidase  
PT levels or other protein or nucleic acid that reduces active cytokinins in  
XX a plant.  
XX Claim 3; Page 168-169; 177pp; English.

XX The present invention relates to a method for stimulating root growth or

CC enhancing the formation of lateral or adventitious roots or altering root  
CC geotropism, which comprises increasing in a plant or plant part the level  
CC of a plant cytokinin oxidase or other protein that reduces the level of  
CC active cytokinins in a plant or plant part. Cytokinin oxidase protein and  
CC coding sequences from Arabidopsis thaliana are also provided. The method  
CC is useful in modifying plant morphological, biochemical and physiological  
CC properties, such as in modifying the initiation, stimulation or  
CC enhancement of root growth, adventitious root formation, lateral root  
CC formation, root geotropism, shoot growth, apical dominance, branching,  
CC timing of senescence, timing of flowering, flower formation, seed  
CC development and/or seed yield. The present sequence is a coding sequence  
CC shown in the invention  
XX

SQ Sequence 1572 BP; 461 A; 327 C; 348 G; 436 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 541 Length: 1572  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-10-014-101B-39 (1-7) x ACC85295 (1-1572)

QY 1 ValGlyGlyThrLeuSerAsn 7  
|||||  
DB 517 GTCGGTGGGACGTTATCAAC 537

RESULT 15

ADH61268  
ID ADH61268 standard; DNA; 1572 BP.

XX ADH61268;

AC 25-MAR-2004 (first entry)

DT Arabidopsis thaliana cytokinin oxidase (CKX1) DNA #2.

DE Male-sterility; cytokinin oxidase; transgenic plant; seed production;

XX CKX1; mouse-ear cross; gene; ds.

XX Arabidopsis thaliana.

XX US2003163847-A1.

XX 28-AUG-2003.

XX 20-DEC-2002; 2002US-00326184.

XX 20-DEC-2001; 2001US-0343129P.

XX (PHAA ) PHARMACIA CORP.

XX Huang S, Crossland LD, Cheikh N, Morris RO;

XX WPI; 2003-897983/82.

XX GENBANK; AF303979.

XX Producing plants characterized by reversible male-sterility, useful for

PT maintaining male sterility in plants, by transforming a plant cell with a

PT nucleic acid construct containing a polynucleotide encoding a cytokinin

PT oxidase.

XX Claim 42; SEQ ID NO 9; 33pp; English.

XX The invention relates to a method for producing a plant characterised by  
 CC reversible male-sterility which involves transforming a plant cell with a  
 CC nucleic acid construct containing a polynucleotide encoding a cytokinin  
 CC oxidase. The method is useful for producing reversible male-sterility in  
 CC transgenic plants, or for maintaining male sterility in plants. The  
 CC method reduces the expense of seed production for existing hybrid plants  
 CC such as corn, but also makes it possible to produce hybrid varieties of  
 CC traditionally non-hybrid crops. The method is also useful for introducing  
 CC economically valuable traits from plants having undesirable production  
 CC characteristics into plants having desirable characteristics. The present  
 CC sequence is Arabidopsis thaliana cytokinin oxidase (CKX1) DNA. This  
 CC sequence is used to illustrate the method of the invention.

XX SQ Sequence 1572 BP; 461 A; 327 C; 348 G; 436 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 541 Length: 1572  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 10 Gaps: 0

US-10-014-101B-39 (1-7) x ADH61268 (1-1572)

Qy 1 ValGlyGlyThrIeuSerAsn 7  
 Db 517 GTCGGTGGGACGTATATCAAC 537

RESULT 16

RAC42983

ID RAC42983 standard; DNA; 1575 BP.

XX AC AAC42983;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 37568.

XX Hybridisation assay; genetic mapping; gene expression control;  
 KW protein identification; signal transduction pathway; metabolic pathway;  
 KW promoter; termination sequence; ss.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

PF 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 16-APR-1999; 99US-0128714P.

PR 19-APR-1999; 99US-0129845P.

PR 21-APR-1999; 99US-0130077P.

PR 23-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 28-APR-1999; 99US-0130891P.

PR 30-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR 04-MAY-1999; 99US-0132484P.

PR 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 06-MAY-1999; 99US-0132487P.

PR 07-MAY-1999; 99US-0132863P.  
 PR 11-MAY-1999; 99US-0134256P.  
 PR 14-MAY-1999; 99US-0134218P.  
 PR 14-MAY-1999; 99US-0134219P.  
 PR 14-MAY-1999; 99US-0134221P.  
 PR 14-MAY-1999; 99US-0134370P.  
 PR 18-MAY-1999; 99US-0134768P.  
 PR 19-MAY-1999; 99US-0134941P.  
 PR 20-MAY-1999; 99US-0135124P.  
 PR 21-MAY-1999; 99US-0135353P.  
 PR 24-MAY-1999; 99US-0135629P.  
 PR 25-MAY-1999; 99US-0136021P.  
 PR 27-MAY-1999; 99US-0136392P.  
 PR 28-MAY-1999; 99US-0136782P.  
 PR 01-JUN-1999; 99US-0137222P.  
 PR 03-JUN-1999; 99US-0137528P.  
 PR 04-JUN-1999; 99US-0137502P.  
 PR 07-JUN-1999; 99US-0137724P.  
 PR 08-JUN-1999; 99US-0138094P.  
 PR 10-JUN-1999; 99US-0138540P.  
 PR 10-JUN-1999; 99US-0138847P.  
 PR 14-JUN-1999; 99US-0139119P.  
 PR 16-JUN-1999; 99US-0139452P.  
 PR 16-JUN-1999; 99US-0139453P.  
 PR 17-JUN-1999; 99US-0139492P.  
 PR 18-JUN-1999; 99US-0139454P.  
 PR 18-JUN-1999; 99US-0139455P.  
 PR 18-JUN-1999; 99US-0139456P.  
 PR 18-JUN-1999; 99US-0139457P.  
 PR 18-JUN-1999; 99US-0139458P.  
 PR 18-JUN-1999; 99US-0139459P.  
 PR 18-JUN-1999; 99US-0139460P.  
 PR 18-JUN-1999; 99US-0139461P.  
 PR 18-JUN-1999; 99US-0139462P.  
 PR 18-JUN-1999; 99US-0139463P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 18-JUN-1999; 99US-0139763P.  
 PR 21-JUN-1999; 99US-0139817P.  
 PR 22-JUN-1999; 99US-0139899P.  
 PR 23-JUN-1999; 99US-0140353P.  
 PR 23-JUN-1999; 99US-0140354P.  
 PR 24-JUN-1999; 99US-0140695P.  
 PR 28-JUN-1999; 99US-0140823P.  
 PR 29-JUN-1999; 99US-0140991P.  
 PR 30-JUN-1999; 99US-0141287P.  
 PR 01-JUL-1999; 99US-0141842P.  
 PR 01-JUL-1999; 99US-0142154P.  
 PR 02-JUL-1999; 99US-0142055P.  
 PR 06-JUL-1999; 99US-0142390P.  
 PR 08-JUL-1999; 99US-0142803P.  
 PR 09-JUL-1999; 99US-0142920P.  
 PR 12-JUL-1999; 99US-0142977P.  
 PR 13-JUL-1999; 99US-0143542P.  
 PR 14-JUL-1999; 99US-0143624P.  
 PR 15-JUL-1999; 99US-0144005P.  
 PR 16-JUL-1999; 99US-0144085P.  
 PR 16-JUL-1999; 99US-0144086P.  
 PR 19-JUL-1999; 99US-0144325P.  
 PR 19-JUL-1999; 99US-0144331P.  
 PR 19-JUL-1999; 99US-0144332P.  
 PR 19-JUL-1999; 99US-0144333P.  
 PR 19-JUL-1999; 99US-0144334P.  
 PR 19-JUL-1999; 99US-0144335P.  
 PR 20-JUL-1999; 99US-0144352P.  
 PR 20-JUL-1999; 99US-0144632P.  
 PR 20-JUL-1999; 99US-0144684P.  
 PR 21-JUL-1999; 99US-0144814P.  
 PR 21-JUL-1999; 99US-0145086P.  
 PR 21-JUL-1999; 99US-0145088P.  
 PR 22-JUL-1999; 99US-0145085P.  
 PR 22-JUL-1999; 99US-0145087P.  
 PR 22-JUL-1999; 99US-0145089P.  
 PR 22-JUL-1999; 99US-0145192P.

PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149262P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150894P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 30-AUG-1999; 99US-0151080P.  
PR 31-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151910P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.

PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores: 542 Length: 1575  
Pred. No.: 35.00 Matches: 7  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 100.00% Gaps: 0  
DB: 3

US-10-014-101b-39 (1-7) x AAC42983 (1-1575)

QY 1 ValGlyGlyThrieuSerAsn 7  
DB 535 GTCGGCGGACGTTGCGAAC 555

RESULT 17  
ABZ14284  
ID ABZ14284 standard; DNA; 1575 BP.  
XX  
AC ABZ14284;  
XX  
XX  
DT 21-JAN-2003 (first entry)  
XX  
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 2089.  
XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.  
KW Arabidopsis thaliana.  
OS Arabidopsis thaliana.  
XX  
PN WO200216655-A2.  
XX  
PD 28-FEB-2002.  
XX  
XX  
PF 24-AUG-2001; 2001WO-US026685.  
XX  
PR 24-AUG-2000; 2000US-0227866P.  
PR 26-JAN-2001; 2001US-0264647P.  
PR 22-JUN-2001; 2001US-0300111P.  
XX  
XX  
PA (SCRI ) SCRIPPS RES INST.  
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX  
XX  
PI Harper JF, Kreps J, Wang X, Zhu T;  
XX  
XX  
DR WPI; 2002-304127/34.  
XX  
PT Identifying a stress condition to which a plant cell has been exposed and  
XX producing plants with increased tolerance to these abiotic stresses.  
PS Claim 144; SEQ ID NO 2089; 577pp + Sequence Listing; English.  
XX  
XX  
CC The invention relates to identifying a stress condition to which a plant  
CC cell has been exposed, comprising: (a) contacting nucleic acid  
CC representative of expressed polynucleotides in the plant cell with an  
CC array or probes representative of the plant cell genome; and (b)  
CC detecting a profile of expressed polynucleotides in the plant cell  
CC characteristic of a stress response. The method is useful in the  
CC production of transgenic plants, cells and seeds and in producing plants  
CC with increased tolerance to abiotic stress. The present sequence is that  
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used

The invention relates to an isolated polynucleotide (I) encoding a novel plant protein (II) having cytokinin oxidase activity. (I) is useful for production of transgenic plants, plant cells or tissues; for production of altered plants, plant cell or tissues; and for effecting the expression of (II) where (I) is operably linked to one or more control sequences. The methods further comprises regenerating a plant from the plant cell. (I) and (II) are useful for stimulating root growth; enhancing the formation of lateral or adventitious roots; altering root geotropism, leading to an increase in yield; and for screening growth promoting chemical of herbicides. (I) is useful for increasing the size of the root meristem; increasing root size; increasing leaf senescence; shoot meristem; delaying leaf senescence and altering leaf senescence; increasing leaf thickness; reducing or increasing the vessel size; inducing parthenocary; improving standability of the seedlings;

Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0

US-10-014-101B-39 (1-7) x ADA68648 (1-1575)

QY 1 ValGlyGlyThrLeuSerAsn 7

DB 535 GTCGGCGGACGTTGTGGAAC 555

RESULT 20

ACC85296

ID ACC85296 standard; cDNA; 1575 BP.

XX

AC ACC85296;

XX

DT 18-SEP-2003 (first entry)

XX

DE Arabidopsis cytokinin oxidase-like protein 4 cDNA.

XX

DE Maize; root growth; root geotropism; cytokinin oxidase; seed size;  
KW embryo size; cotyledon size; transgenic plant; herbicide; gene; ss.

XX

OS Arabidopsis thaliana.

XX

PN WO2003050287-A2.

XX

PD 19-JUN-2003.

XX

PF 10-DEC-2002; 2002WO-EP013990.

XX

PR 10-DEC-2001; 2001US-00014101.

XX

PA (SCHM/) SCHMULLING T.

XX

PA (WERN/) WERNER T.

XX

PI Schmulling T, Werner T;

XX

DR WPI; 2003-541577/51.

XX

PT Stimulating root growth, enhancing lateral or adventitious root formation  
PT or altering root geotropism comprises increasing plant cytokinin oxidase  
PT levels or other protein or nucleic acid that reduces active cytokinins in  
PT a plant.

XX

PS Claim 2; Page 169-170; 177pp; English.

XX

CC The present invention relates to a method for stimulating root growth or  
CC enhancing the formation of lateral or adventitious roots or altering root  
CC geotropism, which comprises increasing in a plant or plant part the level  
CC of a plant cytokinin oxidase or other protein that reduces the level of  
CC active cytokinins in a plant or plant part. Cytokinin oxidase protein and  
CC coding sequences from Arabidopsis thaliana are also provided. The method  
CC is useful in modifying plant morphological, biochemical and physiological  
CC properties, such as in modifying the initiation, stimulation or  
CC enhancement of root growth, adventitious root formation, lateral root  
CC formation, root geotropism, shoot growth, apical dominance, branching,  
CC timing of senescence, timing of flowering, flower formation, seed  
CC development and/or seed yield. The present sequence is a coding sequence  
CC shown in the invention

XX

SQ Sequence 1575 BP; 418 A; 368 C; 349 G; 440 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 542 Length: 1575  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-10-014-101B-39 (1-7) x ACC85296 (1-1575)

QY 1 ValGlyGlyThrLeuSerAsn 7

DB 535 GTCGGCGGACGTTGTGGAAC 555

RESULT 21

ADH61269

ID ADH61269 standard; DNA; 1575 BP.

XX

AC ADH61269;

XX

DT 25-MAR-2004 (first entry)

XX

DE Arabidopsis thaliana cytokinin oxidase (CKX1) DNA #3.

XX

KW Male-sterility; cytokinin oxidase; transgenic plant; seed production;  
KW CKX1; mouse-ear cress; gene; ds.

XX

OS Arabidopsis thaliana.

XX

PN US2003163847-A1.

XX

PD 28-AUG-2003.

XX

PF 20-DEC-2002; 2002US-00326184.

XX

PR 20-DEC-2001; 2001US-0343129P.

XX

PA (PHAA ) PHARMACIA CORP.

XX

PI Huang S, Crossland LD, Cheikh N, Morris RO;

XX

DR WPI; 2003-897983/82.

XX

DR GENBANK; AF303980.

XX

PT Producing plants characterized by reversible male-sterility, useful for  
PT maintaining male sterility in plants, by transforming a plant cell with a  
PT nucleic acid construct containing a polynucleotide encoding a cytokinin  
PT oxidase.

XX

PS Claim 43; SEQ ID NO 10; 33pp; English.

XX

CC The invention relates to a method for producing a plant characterised by  
CC reversible male-sterility which involves transforming a plant cell with a  
CC nucleic acid construct containing a polynucleotide encoding a cytokinin  
CC oxidase. The method is useful for producing reversible male-sterility in  
CC transgenic plants, or for maintaining male sterility in plants. The  
CC method reduces the expense of seed production for existing hybrid plants  
CC such as corn, but also makes it possible to produce hybrid varieties of  
CC traditionally non-hybrid crops. The method is also useful for introducing  
CC economically valuable traits from plants having undesirable production  
CC characteristics into plants having desirable characteristics. The present  
CC sequence is Arabidopsis thaliana cytokinin oxidase (CKX1) DNA. This  
CC sequence is used to illustrate the method of the invention.

XX

SQ Sequence 1575 BP; 418 A; 367 C; 349 G; 441 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 542 Length: 1575  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-10-014-101B-39 (1-7) x ADH61269 (1-1575)

QY 1 ValGlyGlyThrLeuSerAsn 7

DB 535 GTCGGCGGACGTTGTGGAAC 555

RESULT 22

ADH61270  
 ID ADH61270 standard; DNA; 1575 BP.  
 XX  
 AC ADH61270;  
 XX  
 DT 25-MAR-2004 (first entry)  
 XX  
 DE Arabidopsis thaliana cytokinin oxidase (CKX1) DNA #4.  
 XX  
 KW Male-sterility; cytokinin oxidase; transgenic plant; seed production;  
 XX CKX1; mouse-ear cross; gene; ds.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 XX US2003163847-A1.  
 XX  
 XX 28-AUG-2003.  
 XX  
 XX 20-DEC-2002; 2002US-00326184.  
 XX  
 XX 20-DEC-2001; 2001US-0343129P.  
 XX  
 XX (PHAA ) PHARMACIA CORP.  
 XX  
 XX Huang S, Crossland LD, Cheikh N, Morris RO;  
 XX  
 XX WPI; 2003-897983/82.  
 XX GENBANK; AF303981.  
 XX  
 XX Producing plants characterized by reversible male-sterility, useful for  
 PT maintaining male sterility in plants, by transforming a plant cell with a  
 PT nucleic acid construct containing a polynucleotide encoding a cytokinin  
 PT oxidase.  
 XX  
 XX Claim 44; SEQ ID NO 11; 33pp; English.  
 XX  
 CC The invention relates to a method for producing a plant characterised by  
 CC reversible male-sterility which involves transforming a plant cell with a  
 CC nucleic acid construct containing a polynucleotide encoding a cytokinin  
 CC oxidase. The method is useful for producing reversible male-sterility in  
 CC transgenic plants, or for maintaining male sterility in plants. The  
 CC method reduces the expense of seed production for existing hybrid plants  
 CC such as corn, but also makes it possible to produce hybrid varieties of  
 CC traditionally non-hybrid crops. The method is also useful for introducing  
 CC economically valuable traits from plants having undesirable production  
 CC characteristics into plants having desirable characteristics. The present  
 CC sequence is Arabidopsis thaliana cytokinin oxidase (CKX1) DNA. This  
 CC sequence is used to illustrate the method of the invention.  
 XX  
 XX Sequence 1575 BP; 390 A; 328 C; 451 G; 406 T; 0 U; 0 Other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 542 Length: 1575  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 10 Gaps: 0  
 US-10-014-101B-39 (1-7) x ADH61270 (1-1575)  
 QY 1 ValGlyGlyThrLeuSerAsn 7  
 DB 502 GTGGAGGTACGTTGTCAAT 522  
 RESULT 23  
 ADA69574  
 ID ADA69574 standard; DNA; 1587 BP.  
 XX  
 AC ADA69574;  
 XX  
 XX 20-NOV-2003 (first entry)  
 XX  
 XX Arabidopsis thaliana cytokinin oxidase (CKX1) DNA #4.  
 XX  
 XX US2003163847-A1.  
 XX  
 XX 28-AUG-2003.  
 XX  
 XX 20-DEC-2002; 2002US-00326184.  
 XX  
 XX 20-DEC-2001; 2001US-0343129P.  
 XX  
 XX (PHAA ) PHARMACIA CORP.  
 XX  
 XX Huang S, Crossland LD, Cheikh N, Morris RO;  
 XX  
 XX WPI; 2003-897983/82.  
 XX GENBANK; AF303981.  
 XX  
 XX Producing plants characterized by reversible male-sterility, useful for  
 PT maintaining male sterility in plants, by transforming a plant cell with a  
 PT nucleic acid construct containing a polynucleotide encoding a cytokinin  
 PT oxidase.  
 XX  
 XX Claim 44; SEQ ID NO 11; 33pp; English.  
 XX  
 CC The invention relates to a method for producing a plant characterised by  
 CC reversible male-sterility which involves transforming a plant cell with a  
 CC nucleic acid construct containing a polynucleotide encoding a cytokinin  
 CC oxidase. The method is useful for producing reversible male-sterility in  
 CC transgenic plants, or for maintaining male sterility in plants. The  
 CC method reduces the expense of seed production for existing hybrid plants  
 CC such as corn, but also makes it possible to produce hybrid varieties of  
 CC traditionally non-hybrid crops. The method is also useful for introducing  
 CC economically valuable traits from plants having undesirable production  
 CC characteristics into plants having desirable characteristics. The present  
 CC sequence is Arabidopsis thaliana cytokinin oxidase (CKX1) DNA. This  
 CC sequence is used to illustrate the method of the invention.  
 XX  
 XX Sequence 1575 BP; 390 A; 328 C; 451 G; 406 T; 0 U; 0 Other;  
 SQ

DE Rice gene, SEQ ID 2897.  
 XX  
 KW Plant; bacterial infection; fungal infection; viral infection; rice;  
 KW gene; ds.  
 XX  
 OS Oryza sativa.  
 XX  
 XX WO2003000898-A1.  
 XX  
 XX 03-JAN-2003.  
 XX  
 XX 22-JUN-2001; 2001WO-IB001105.  
 XX  
 XX 22-JUN-2001; 2001WO-IB001105.  
 XX  
 XX (SYGN ) SYNGENTA PARTICIPATIONS AG.  
 XX  
 XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
 XX Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
 XX WPI; 2003-175290/17.  
 XX  
 XX Identifying at least one gene involved in plant resistance or response to  
 PT pathogenic infection for conferring resistance or tolerance to a plant to  
 PT bacterial, fungal or viral infection by determining or detecting plant  
 PT gene expression.  
 XX  
 XX Claim 6; SEQ ID NO 2897; 899pp; English.  
 XX  
 CC The present invention relates to a method (M1) for identifying genes  
 CC involved in plant resistance or response to pathogenic infection. M1  
 CC comprises identifying a gene whose expression is significantly altered in  
 CC the incompatible interaction of plant gene expression relative to  
 CC expression of the gene in an uninfected plant, in a mutant plant that  
 CC does not express a gene associated with response to pathogenic infection,  
 CC or in a corresponding incompatible or compatible interaction. (M1) is  
 CC useful for conferring resistance to resistance or tolerance to a plant to  
 CC bacterial, fungal or viral infection. The present sequence was used to  
 CC illustrate the invention.  
 XX  
 XX Sequence 1587 BP; 307 A; 462 C; 458 G; 356 T; 0 U; 4 Other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 547 Length: 1587  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 8 Gaps: 0  
 US-10-014-101B-39 (1-7) x ADA69574 (1-1587)  
 QY 1 ValGlyGlyThrLeuSerAsn 7  
 DB 523 GTCGGTGGCAGCGCTCTCAAC 543  
 RESULT 24  
 ADA69773  
 ID ADA69773 standard; DNA; 1590 BP.  
 XX  
 AC ADA69773;  
 XX  
 XX 20-NOV-2003 (first entry)  
 XX  
 XX Rice gene, SEQ ID 3096.  
 XX  
 XX Plant; bacterial infection; fungal infection; viral infection; rice;  
 KW gene; ds.  
 XX  
 OS Oryza sativa.  
 XX  
 XX WO2003000898-A1.  
 XX

```

PD 03-JAN-2003.
XX
PF 22-JUN-2001; 2001WO-IB001105.
XX
PR 22-JUN-2001; 2001WO-IB001105.
XX
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX
XX WPI; 2003-175290/17.
XX
XX Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
PS Claim 6; SEQ ID NO 3096; 899pp; English.
XX
CC The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
SQ Sequence 1590 BP; 356 A; 465 C; 460 G; 309 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 548 Length: 1590
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0

US-10-014-101B-39 (1-7) x ADA69773 (1-1590)

Qy 1 ValGlyGlyThrIeuSerAsn 7
Db 520 GTCGGTGGCACCTTGTGGAAT 540

RESULT 25
ADA69430
ID ADA69430 standard; DNA; 1593 BP.
XX
AC ADA69430;
XX
XX 20-NOV-2003 (first entry)
DT
DE Rice gene, SEQ ID 2753.
XX
XX Plant; bacterial infection; fungal infection; viral infection; rice;
XX gene; ds.
XX
OS Oryza sativa.
XX
XX WO2003000898-A1.
XX
XX 03-JAN-2003.
XX
XX 22-JUN-2001; 2001WO-IB001105.
XX
XX 22-JUN-2001; 2001WO-IB001105.
XX
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI

```

---

```

PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX
XX WPI; 2003-175290/17.
XX
XX Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
PS Claim 6; SEQ ID NO 2753; 899pp; English.
XX
CC The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
SQ Sequence 1593 BP; 232 A; 523 C; 569 G; 269 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 549 Length: 1593
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0

US-10-014-101B-39 (1-7) x ADA69430 (1-1593)

Qy 1 ValGlyGlyThrIeuSerAsn 7
Db 517 GTCGGCGGCACCTCTCCAAC 537

RESULT 26
AAx02914
ID AAx02914 standard; DNA; 1605 BP.
XX
AC AAx02914;
XX
XX 20-MAY-1999 (first entry)
DT
DE Z. mays cck1 DNA coding region.
XX
XX Cytokinin oxidase; cck1; transgenic plant; altered growth behaviour;
XX cytokinin-associated pathogenesis; resistance; fungi; nematode; assay;
XX grain yield; secondary growth; metabolism; senescence; ss.
XX
XX Zea mays.
XX
XX WO9906571-A1.
XX
XX 11-FEB-1999.
XX
XX 30-JUL-1998; 98WO-US015844.
XX
XX 30-JUL-1997; 97US-0054268P.
XX
XX 29-JUL-1998; 98US-00124541.
XX
XX (UMOR ) UNIV MISSOURI.
XX
XX Morris RO;
XX
XX WPI; 1999-153800/13.
XX
XX P-PSDB; AAW93007.
XX
XX New cytokinin oxidase from maize - used to generate transgenic plants
XX with, e.g. better disease resistance and growth characteristics.
XX

```

PS Claim 7b; Page 64-69; 140pp; English.

XX This sequence encodes a Zea mays cytokinin oxidase, cckxl. This gene is used to generate transgenic plants in which cytokinin-associated pathogenesis or growth behaviour is altered. Particularly applications include generation of plants with increased resistance to fungi and nematodes, increased grain yield and superior secondary growth properties. Host cells are used for production of recombinant cckxl nucleic acid which is useful in an assay for determining cytokinin concentrations, and for its studying effects on plant growth and metabolism, including senescence

XX SQ Sequence 1605 BP; 236 A; 560 C; 561 G; 248 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	553	Length:	1605
Score:	35.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-10-014-101B-39 (1-7) x AAX02914 (1-1605)

QY 1 ValGlyGlyThrLeuSerAsn 7  
|||||

Db 523 GTCGGCGGCACGCTGTCCAAC 543

RESULT 27

AAC86501

ID AAC86501 standard; DNA; 1608 BP.

AC AAC86501;

DT 19-MAR-2001 (first entry)

XX DNA encoding a maize cytokinin oxidase polypeptide.

XX Cytokinin oxidase; temporal gene expression; spatial gene expression; plant seed; cytokinin modulating gene; transgenic plant; seed size; stress tolerance; yield stability; tip kernel abortion; seed set; ss.

XX Zea mays.

XX Key Location/Qualifiers

FT 1. .1608

FT /\*tag= a

FT /product= "cytokinin oxidase"

XX WO200063401-A1.

PN 13-APR-2000; 2000WO-US009943.

XX 26-OCT-2000.

PF 16-APR-1999; 99US-0129844P.

XX (PION-) PIONEER HI-BRED INT INC.

PA Habben JE, Zinselmeier C, Tomes D;

XX WPI; 2000-672743/65.

DR P-PSDB; AAB30691.

XX Novel recombinant DNA construct useful for producing transgenic plants having enhanced levels of cytokinin expression, improved stress tolerance and yield stability.

PS Disclosure; Page 63-65; 76pp; English.

XX The present sequence encodes a maize cytokinin oxidase polypeptide. It is used to produce the recombinant DNA molecules of the invention. These comprise a genetic construct consisting of a promoter directing temporal

CC and/or spatial gene expression in plant seed operatively linked to a cytokinin modulating gene. The recombinant DNA molecules are useful for producing fertile, transgenic plants capable of regulated expression of a cytokinin modulating gene in developing seeds. They are also useful for improving stress tolerance and yield stability in plants. The preferential expression of recombinant DNA molecules of the invention occurs about 14-25 days after pollination. The transgenic plants thus produced have enhanced levels of cytokinin expression exhibit improved seed size, decreased tip kernel abortion and increased seed set during unfavourable environmental conditions

XX SQ Sequence 1608 BP; 237 A; 561 C; 560 G; 250 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	555	Length:	1608
Score:	35.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	3	Gaps:	0

US-10-014-101B-39 (1-7) x AAC86501 (1-1608)

QY 1 ValGlyGlyThrLeuSerAsn 7  
|||||

Db 529 GTCGGCGGCACGCTGTCCAAC 549

RESULT 28

ABK28628

ID ABK28628 standard; cDNA; 1611 BP.

XX AC ABK28628;

DT 09-APR-2002 (first entry)

XX cDNA encoding A. thaliana cytokinin oxidase AtCKX5.

XX Arabidopsis thaliana; plant; cytokinin oxidase; transgenic plant; root growth; lateral root; adventitious root; root geotropism; herbicide; root meristem; shoot meristem; leaf senescence; parthenocarp; gene; ss.

XX Arabidopsis thaliana.

OS WO200196580-A2.

PN 20-DEC-2001.

XX 18-JUN-2001; 2001WO-EP006833.

XX 16-JUN-2000; 2000EP-00870132.

PR 27-DEC-2000; 2000US-0258415P.

PR 16-MAR-2001; 2001EP-00870053.

XX (SCHM/) SCHMULLING T.

PA (WERN/) WERNER T.

XX Schmullling T, Werner T;

PI WPI; 2002-130736/17.

XX Polynucleotide encoding novel plant protein having cytokinin oxidase activity and the protein useful for stimulating root growth, enhancing the formation of lateral or adventitious roots, altering root geotropism.

PS Claim 3; Page 148-149; 154pp; English.

XX The invention relates to an isolated polynucleotide (I) encoding a novel plant protein (II) having cytokinin oxidase activity. (I) is useful for production of transgenic plants, plant cells or tissues; for production of altered plants, plant cell or tissues; and for effecting the expression of (II) where (I) is operably linked to one or more control sequences. The methods further comprises regenerating a plant from the plant cell. (I) and (II) are useful for stimulating root growth;

CC enhancing the formation of lateral or adventitious roots; altering root  
 CC geotropism, leading to an increase in yield; and for screening growth  
 CC promoting chemical of herbicides. (I) is useful for increasing the size  
 CC of the root meristem; increasing root size; increasing the size of the  
 CC shoot meristem; delaying leaf senescence and altering leaf senescence;  
 CC increasing leaf thickness; reducing or increasing the vessel size;  
 CC inducing parthenocarp; improving standability of the seedlings;  
 CC increasing branching and for improving lodging resistance. Antibody (III)  
 CC to (II) is useful for identifying and obtaining proteins interacting with  
 CC (II) comprising a screening assay, preferably a two-hybrid screening  
 CC assay. ABK28632-ABK28633 represent Arabidopsis thaliana cytokinin oxidase  
 CC coding sequences and PCR primers of the invention

XX SQ Sequence 1611 BP; 416 A; 379 C; 409 G; 407 T; 0 U; 0 Other;

Alignment Scores: Pred. No.: 556 Length: 1611  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-10-014-101B-39 (1-7) x ABK28628 (1-1611)

Qy 1 ValGlyGlyThrLeuSerAsn 7  
 Db 499 GTTGAGGTACACTCTCCAAT 519

RESULT 29

ACC85297 ID ACC85297 standard; cDNA; 1611 BP.

XX AC ACC85297;

XX DT 18-SEP-2003 (first entry)

XX DE Arabidopsis cytokinin oxidase-like protein 5 cDNA #1.

XX Maize; root growth; root geotropism; cytokinin oxidase; seed size;  
 XX embryo size; cotyledon size; transgenic plant; herbicide; gene; ss.

XX Arabidopsis thaliana.

XX WO2003050287-A2.

XX 19-JUN-2003.

XX 10-DEC-2002; 2002WO-EP013990.

XX 10-DEC-2001; 2001US-00014101.

XX (SCHM/) SCHMULLING T.  
 XX (WERN/) WERNER T.

XX Schnulling T, Werner T;

XX WPI; 2003-541577/51.

XX Stimulating root growth, enhancing lateral or adventitious root formation  
 XX or altering root geotropism comprises increasing plant cytokinin oxidase  
 XX levels or other protein or nucleic acid that reduces active cytokinins in  
 XX a plant.

XX Claim 3; Page 170-171; 177pp; English.

XX The present invention relates to a method for stimulating root growth or  
 XX enhancing the formation of lateral or adventitious roots or altering root  
 XX geotropism, which comprises increasing in a plant or plant part the level  
 XX of a plant cytokinin oxidase or other protein that reduces the level of  
 XX active cytokinins in a plant or plant part. Cytokinin oxidase protein and  
 XX coding sequences from Arabidopsis thaliana are also provided. The method  
 XX is useful in modifying plant morphological, biochemical and physiological

CC properties, such as in modifying the initiation, stimulation or  
 CC enhancement of root growth, adventitious root formation, lateral root  
 CC formation, root geotropism, shoot growth, apical dominance, branching,  
 CC timing of senescence, timing of flowering, flower formation, seed  
 CC development and/or seed yield. The present sequence is a coding sequence  
 CC shown in the invention

XX SQ Sequence 1611 BP; 416 A; 379 C; 409 G; 407 T; 0 U; 0 Other;

Alignment Scores: Pred. No.: 556 Length: 1611  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-10-014-101B-39 (1-7) x ACC85297 (1-1611)

Qy 1 ValGlyGlyThrLeuSerAsn 7  
 Db 499 GTTGAGGTACACTCTCCAAT 519

RESULT 30

ABK28632 ID ABK28632 standard; cDNA; 1620 BP.

XX AC ABK28632;

XX DT 09-APR-2002 (first entry)

XX DE cDNA encoding A. thaliana cytokinin oxidase AtCKX5 (long version).

XX Arabidopsis thaliana; plant; cytokinin oxidase; transgenic plant;  
 XX root growth; lateral root; adventitious root; root geotropism; herbicide;  
 XX root meristem; shoot meristem; leaf senescence; parthenocarp; gene; ss.

XX Arabidopsis thaliana.

XX WO200196580-A2.

XX 20-DEC-2001.

XX 18-JUN-2001; 2001WO-EP006833.

XX 16-JUN-2000; 2000EP-00870132.

XX 27-DEC-2000; 2000US-0258415P.

XX 16-MAR-2001; 2001EP-00870053.

XX (SCHM/) SCHMULLING T.

XX (WERN/) WERNER T.

XX Schnulling T, Werner T;

XX WPI; 2002-130736/17.

XX P-PSDB; AA081974.

XX Polynucleotide encoding novel plant protein having cytokinin oxidase  
 XX activity and the protein useful for stimulating root growth, enhancing  
 XX the formation of lateral or adventitious roots, altering root geotropism.

XX Claim 3; Page 151; 154pp; English.

XX The invention relates to an isolated polynucleotide (I) encoding a novel  
 XX plant protein (II) having cytokinin oxidase activity. (I) is useful for  
 XX production of transgenic plants, plant cells or tissues; for production  
 XX of altered plants, plant cell or tissues; and for effecting the  
 XX expression of (II) where (I) is operably linked to one or more control  
 XX sequences. The methods further comprises regenerating a plant from the  
 XX plant cell. (I) and (II) are useful for stimulating root growth;  
 XX enhancing the formation of lateral or adventitious roots; altering root  
 XX geotropism, leading to an increase in yield; and for screening growth  
 XX promoting chemical of herbicides. (I) is useful for increasing the size

CC of the root meristem; increasing root size; increasing the size of the  
 CC shoot meristem; delaying leaf senescence and altering leaf senescence;  
 CC increasing leaf thickness; reducing or increasing the vessel size;  
 CC inducing parthenocarp; improving standability of the seedlings;  
 CC increasing branching and for improving lodging resistance. Antibody (III)  
 CC to (II) is useful for identifying and obtaining proteins interacting with  
 CC (II) comprising a screening assay, preferably a two-hybrid screening  
 CC assay. ABK28606-ABK28633 represent Arabidopsis thaliana cytokinin oxidase  
 CC coding sequences and PCR primers of the invention  
 XX  
 SQ Sequence 1620 BP; 419 A; 380 C; 411 G; 410 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 559 Length: 1620  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-10-014-101B-39 (1-7) x ABK28632 (1-1620)

OY 1 ValGlyGlyThrLeuSerAsn 7  
 |||||  
 DB 508 GTTGAGGTACACTCTCCAAT 528

## RESULT 31

ACC85301  
 ID ACC85301 standard; cDNA; 1620 BP.

AC ACC85301;

DT 18-SEP-2003 (first entry)

XX Arabidopsis cytokinin oxidase-like protein 5 cDNA #2.

XX Maize; root growth; root geotropism; cytokinin oxidase; seed size;  
 KW embryo size; cotyledon size; transgenic plant; herbicide; gene; ss.

XX Arabidopsis thaliana.

XX WO2003050287-A2.

XX 19-JUN-2003.

XX 10-DEC-2002; 2002WO-EP013990.

XX 10-DEC-2001; 2001US-00014101.

XX (SCHM/) SCHMULLING T.

XX (WERN/) WERNER T.

XX Schmulling T, Werner T;

XX WPI; 2003-541577/51.

XX Stimulating root growth, enhancing lateral or adventitious root formation  
 PT or altering root geotropism comprises increasing plant cytokinin oxidase  
 PT levels or other protein or nucleic acid that reduces active cytokinins in  
 PT a plant.

XX Claim 3; Page 174-175; 177pp; English.

XX The present invention relates to a method for stimulating root growth or  
 CC enhancing the formation of lateral or adventitious roots or altering root  
 CC geotropism, which comprises increasing in a plant or plant part the level  
 CC of a plant cytokinin oxidase or other protein that reduces the level of  
 CC active cytokinins in a plant or plant part. Cytokinin oxidase protein and  
 CC coding sequences from Arabidopsis thaliana are also provided. The method  
 CC is useful in modifying plant morphological, biochemical and physiological  
 CC properties, such as in modifying the initiation, stimulation or  
 CC enhancement of root growth, adventitious root formation, lateral root  
 CC formation, root geotropism, shoot growth, apical dominance, branching,

CC timing of senescence, timing of flowering, flower formation, seed  
 CC development and/or seed yield. The present sequence is a coding sequence  
 CC shown in the invention

SQ Sequence 1620 BP; 419 A; 380 C; 411 G; 410 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 559 Length: 1620  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-10-014-101B-39 (1-7) x ACC85301 (1-1620)

OY 1 ValGlyGlyThrLeuSerAsn 7  
 |||||  
 DB 508 GTTGAGGTACACTCTCCAAT 528

## RESULT 32

ADH61271

ID ADH61271 standard; DNA; 1623 BP.

XX AC ADH61271;

XX 25-MAR-2004 (first entry)

XX Arabidopsis thaliana cytokinin oxidase (CKX1) DNA #5.

XX Male-sterility; cytokinin oxidase; transgenic plant; seed production;  
 KW CKX1; mouse-ear cross; gene; ds.

XX Arabidopsis thaliana.

XX US2003163847-A1.

XX 28-AUG-2003.

XX 20-DEC-2002; 2002US-00326184.

XX 20-DEC-2001; 2001US-0343129P.

XX (PHAA ) PHARMACIA CORP.

XX Huang S, Crossland LD, Cheikh N, Morris RO;

XX WPI; 2003-897983/82.

XX GENBANK; AF303982.

XX Producing plants characterized by reversible male-sterility, useful for  
 PT maintaining male sterility in plants, by transforming a plant cell with a  
 PT nucleic acid construct containing a polynucleotide encoding a cytokinin  
 PT oxidase.

XX Claim 45; SEQ ID NO 12; 33pp; English.

XX The invention relates to a method for producing a plant characterised by  
 CC reversible male-sterility which involves transforming a plant cell with a  
 CC nucleic acid construct containing a polynucleotide encoding a cytokinin  
 CC oxidase. The method is useful for producing reversible male-sterility in  
 CC transgenic plants, or for maintaining male sterility in plants. The  
 CC method reduces the expense of seed production for existing hybrid plants  
 CC such as corn, but also makes it possible to produce hybrid varieties of  
 CC traditionally non-hybrid crops. The method is also useful for introducing  
 CC economically valuable traits from plants having undesirable production  
 CC characteristics into plants having desirable characteristics. The present  
 CC sequence is Arabidopsis thaliana cytokinin oxidase (CKX1) DNA. This  
 CC sequence is used to illustrate the method of the invention.

XX Sequence 1623 BP; 421 A; 382 C; 412 G; 408 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 560 Length: 1623  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-10-014-101B-39 (1-7) x ADH61271 (1-1623)

Qy 1 ValGlyGlyThrLeuSerAsn 7

Db 511 GTTGAGGTACTCTCCCAT 531

RESULT 33

ADH61279

ID ADH61279 standard; DNA; 1655 BP.

XX

AC ADH61279;

XX

DT 25-MAR-2004 (first entry)

XX

DE Arabidopsis thaliana cytokinin oxidase (CKX1) DNA #6.

XX

KW Male-sterility; cytokinin oxidase; transgenic plant; seed production;

KW CKX1; mouse-ear cross; gene; ds.

XX

OS Arabidopsis thaliana.

XX

PN US2003163847-A1.

XX

PD 28-AUG-2003.

XX

PF 20-DEC-2002; 2002US-00326184.

XX

PR 20-DEC-2001; 2001US-0343129P.

XX

PA (PHAA ) PHARMACIA CORP.

XX

PI Huang S, Croseland LD, Cheikh N, Morris RO;

XX

DR WPI; 2003-897983/82.

XX

DR GENBANK; BT000179.

XX

PT Producing plants characterized by reversible male-sterility, useful for maintaining male sterility in plants, by transforming a plant cell with a nucleic acid construct containing a polynucleotide encoding a cytokinin oxidase.

XX

PS Disclosure; SEQ ID NO 20; 33pp; English.

XX

CC The invention relates to a method for producing a plant characterized by reversible male-sterility which involves transforming a plant cell with a nucleic acid construct containing a polynucleotide encoding a cytokinin oxidase. The method is useful for producing reversible male-sterility in transgenic plants, or for maintaining male sterility in plants. The method reduces the expense of seed production for existing hybrid plants such as corn, but also makes it possible to produce hybrid varieties of traditionally non-hybrid crops. The method is also useful for introducing economically valuable traits from plants having undesirable production characteristics into plants having desirable characteristics. The present invention is Arabidopsis thaliana cytokinin oxidase (CKX1) DNA. This sequence is used to illustrate the method of the invention.

XX

SQ Sequence 1655 BP; 444 A; 380 C; 358 G; 473 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 573 Length: 1655

Score: 35.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 10 Gaps: 0

US-10-014-101B-39 (1-7) x ADH61279 (1-1655)

Qy 1 ValGlyGlyThrLeuSerAsn 7

Db 535 GTCGGGGACGCTGTCCAAC 555

RESULT 34

ADA70177

ID ADA70177 standard; DNA; 1677 BP.

XX

AC ADA70177;

XX

DT 20-NOV-2003 (first entry)

XX

DE Rice gene, SEQ ID 3500.

XX

KW Plant; bacterial infection; fungal infection; viral infection; rice;

KW gene; ds.

XX

OS Oryza sativa.

XX

PN WO2003000898-A1.

XX

PD 03-JAN-2003.

XX

PF 22-JUN-2001; 2001WO-IB001105.

XX

PR 22-JUN-2001; 2001WO-IB001105.

XX

PA (SYGN ) SYNGENTA PARTICIPATIONS AG.

XX

PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;

XX

PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

XX

DR WPI; 2003-175290/17.

XX

PT Identifying at least one gene involved in plant resistance or response to pathogenic infection for conferring resistance or tolerance to a plant to bacterial, fungal or viral infection by determining or detecting plant gene expression.

XX

PS Claim 6; SEQ ID NO 3500; 899pp; English.

XX

CC The present invention relates to a method (M1) for identifying genes involved in plant resistance or response to pathogenic infection. M1 comprises identifying a gene whose expression is significantly altered in the incompatible interaction of plant gene expression relative to expression of the gene in an uninfected plant, in a mutant plant that does not express a gene associated with response to pathogenic infection, or in a corresponding incompatible or compatible interaction. (M1) is useful for conferring resistance to resistance or tolerance to a plant to bacterial, fungal or viral infection. The present sequence was used to illustrate the invention.

XX

SQ Sequence 1677 BP; 219 A; 608 C; 582 G; 267 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 581 Length: 1677

Score: 35.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 8 Gaps: 0

US-10-014-101B-39 (1-7) x ADA70177 (1-1677)

Qy 1 ValGlyGlyThrLeuSerAsn 7

Db 532 GTCGGGGACGCTGTCCAAC 552

RESULT 35

ADS49453/C

ID ADS49453 standard; cDNA; 1719 BP.

XX ADS49453;  
 AC  
 XX  
 DT 02-DEC-2004 (first entry)  
 XX  
 DE Bacterial polynucleotide #4196.  
 XX  
 KW Recombinant DNA construct; transformed plant; improved plant property;  
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
 KW pathogen tolerance; pest tolerance; plant disease resistance;  
 KW cell cycle pathway modification; plant growth regulator;  
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
 KW bacterial polynucleotide; gene; ss.  
 XX  
 OS Bacteria.  
 XX  
 XX US2003233675-A1.  
 XX  
 XX 18-DEC-2003.  
 PD  
 XX 20-FEB-2003; 2003US-00369493.  
 XX  
 XX 21-FEB-2002; 2002US-0360039P.  
 XX  
 XX (CAOY/) CAO Y.  
 PA (HINK/) HINKLE G J.  
 PA (SLAT/) SLATER S C.  
 PA (CHEN/) CHEN X.  
 PA (GOLD/) GOLDMAN B S.  
 XX  
 XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
 PI WPI; 2004-061375/06.  
 XX  
 XX New recombinant DNA construct comprising a promoter positioned to provide  
 PT for expression of a polynucleotide encoding a polypeptide from a  
 PT microbial source, useful for producing plants with improved properties.  
 XX  
 PS Claim 1; SEQ ID NO 27883; 122pp; English.  
 XX  
 CC The invention relates to a recombinant DNA construct comprising a  
 CC promoter functional in a plant cell, where the promoter is positioned to  
 CC provide for expression of a polynucleotide encoding a polypeptide from a  
 CC microbial source. The invention also relates to a transformed plant  
 CC comprising the recombinant DNA construct and a method of producing a  
 CC transformed plant having an improved property. The plant is a crop plant  
 CC such as maize or soybean. The method of producing a transformed plant  
 CC having an improved property comprises transforming a plant with the  
 CC recombinant DNA construct and growing the transformed plant, where the  
 CC polynucleotide or polypeptide is useful for improving plant properties.  
 CC The recombinant DNA construct is useful for producing plants with  
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
 CC increased resistance to plant disease, better growth rate by modification  
 CC of the cell cycle pathway with plant growth regulators, increased rate of  
 CC homologous recombination, modified seed oil or protein yield and/or  
 CC content, improved yield by modification of carbohydrate, nitrogen or  
 CC phosphorus use and/or uptake, by modification of photosynthesis or by  
 CC providing improved plant growth and development under at least one stress  
 CC condition, improved lignin production or improved galactomannan  
 CC production. This sequence represents a bacterial polynucleotide used in  
 CC the scope of the invention. Note: The sequence data for this patent did  
 CC not form part of the printed specification but was obtained in electronic  
 CC format from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).  
 XX  
 SQ Sequence 1719 BP; 413 A; 466 C; 489 G; 351 T; 0 U; 0 Other;  
 XX  
 Alignment Scores:  
 Pred. No.: 597 Length: 1719  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0  
 DB: 13 Gaps: 0  
 US-10-014-101b-39 (1-7) x ADS49453 (1-1719)  
 QY 1 ValGlyGlyThrLeuSerAsn 7  
 Db 503 GTAGGCGGACGCTCTCAAC 483  
 RESULT 36  
 ABK28624  
 ID ABK28624 standard; cDNA; 1728 BP.  
 XX AC ABK28624;  
 XX  
 DT 09-APR-2002 (first entry)  
 XX  
 DE cDNA encoding A. thaliana cytokinin oxidase AtCKX1.  
 XX  
 KW Arabidopsis thaliana; plant; cytokinin oxidase; transgenic plant;  
 KW root growth; lateral root; adventitious root; root geotropism; herbicide;  
 KW root meristem; shoot meristem; leaf senescence; parthenocarpy; gene; ss.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 XX WO200196580-A2.  
 XX  
 PD 20-DEC-2001.  
 XX  
 XX 18-JUN-2001; 2001WO-EP006833.  
 XX  
 XX 16-JUN-2000; 2000EP-00870132.  
 PR 27-DEC-2000; 2000US-0258415P.  
 PR 16-MAR-2001; 2001EP-00870053.  
 XX  
 XX (SCHW/) SCHMULLING T.  
 PA (WERN/) WERNER T.  
 XX  
 PI Schmullling T, Werner T;  
 XX  
 DR WPI; 2002-130736/17.  
 XX  
 PT Polynucleotide encoding novel plant protein having cytokinin oxidase  
 PT activity and the protein useful for stimulating root growth, enhancing  
 PT the formation of lateral or adventitious roots, altering root geotropism.  
 XX  
 PS Claim 2; Page 146; 154pp; English.  
 XX  
 CC The invention relates to an isolated polynucleotide (I) encoding a novel  
 CC plant protein (II) having cytokinin oxidase activity. (I) is useful for  
 CC production of transgenic plants, plant cells or tissues; for production  
 CC of altered plants, plant cell or tissues; and for effecting the  
 CC expression of (II) where (I) is operably linked to one or more control  
 CC sequences. The methods further comprises regenerating a plant from the  
 CC plant cell. (I) and (II) are useful for stimulating root growth;  
 CC enhancing the formation of lateral or adventitious roots; altering root  
 CC geotropism, leading to an increase in yield; and for screening growth  
 CC promoting chemical of herbicides. (I) is useful for increasing the size  
 CC of the root meristem; increasing root size; increasing the size of the  
 CC shoot meristem; delaying leaf senescence and altering leaf senescence;  
 CC increasing leaf thickness; reducing or increasing the vessel size;  
 CC inducing parthenocarpy; improving standability of the seedlings;  
 CC increasing branching and for improving lodging resistance. Antibody (III)  
 CC to (II) is useful for identifying and obtaining proteins interacting with  
 CC (II) comprising a screening assay, preferably a two-hybrid screening  
 CC assay. ABK28606-ABK28633 represent Arabidopsis thaliana cytokinin oxidase  
 CC coding sequences and PCR primers of the invention  
 XX  
 SQ Sequence 1728 BP; 544 A; 390 C; 332 G; 462 T; 0 U; 0 Other;  
 XX  
 Alignment Scores:  
 Pred. No.: 601 Length: 1728  
 Score: 35.00 Matches: 7

Percent Similarity: 100.00%    Conservative: 0  
 Best Local Similarity: 100.00%    Mismatches: 0  
 Query Match: 100.00%    Indels: 0  
 DB: 6    Gaps: 0

US-10-014-101B-39 (1-7) x ABK28624 (1-1728)

Qy 1 ValGlyGlyThrLeuSerAsn 7  
 |||||  
 Db 574 GTTGGAGGTACACTATCTAAT 594

## RESULT 37

ACC85293  
 ID ACC85293 standard; cDNA; 1728 BP.  
 XX  
 AC ACC85293;  
 XX  
 DT 18-SEP-2003 (first entry)  
 XX  
 DE Arabidopsis cytokinin oxidase-like protein 1 cDNA.  
 XX  
 KW Maize; root growth; root geotropism; cytokinin oxidase; seed size;  
 KW embryo size; cotyledon size; transgenic plant; herbicide; gene; ss.  
 XX  
 OS Arabidopsis thaliana.  
 XX

WO2003050287-A2.

19-JUN-2003.

10-DEC-2002; 2002WO-EP013990.

10-DEC-2001; 2001US-00014101.

(SCHM/) SCHMULLING T.

(WERN/) WERNER T.

Schmulling T, Werner T;

WPI; 2003-541577/51.

Stimulating root growth, enhancing lateral or adventitious root formation or altering root geotropism comprises increasing plant cytokinin oxidase levels or other protein or nucleic acid that reduces active cytokinins in a plant.

Claim 122; Page 166-167; 177pp; English.

The present invention relates to a method for stimulating root growth or enhancing the formation of lateral or adventitious roots or altering root geotropism, which comprises increasing in a plant or plant part the level of a plant cytokinin oxidase or other protein that reduces the level of active cytokinins in a plant or plant part. Cytokinin oxidase protein and coding sequences from Arabidopsis thaliana are also provided. The method is useful in modifying plant morphological, biochemical and physiological properties, such as in modifying the initiation, stimulation or enhancement of root growth, adventitious root formation, lateral root formation, root geotropism, shoot growth, apical dominance, branching, timing of senescence, timing of flowering, flower formation, seed development and/or seed yield. The present sequence is a coding sequence shown in the invention

Sequence 1728 BP; 544 A; 390 C; 332 G; 462 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 601    Length: 1728  
 Score: 35.00    Matches: 7  
 Percent Similarity: 100.00%    Conservative: 0  
 Best Local Similarity: 100.00%    Mismatches: 0  
 Query Match: 100.00%    Indels: 0  
 DB: 9    Gaps: 0

US-10-014-101B-39 (1-7) x ACC85293 (1-1728)

Qy 1 ValGlyGlyThrLeuSerAsn 7  
 |||||  
 Db 574 GTTGGAGGTACACTATCTAAT 594

## RESULT 38

ADO06496  
 ID ADO06496 standard; DNA; 1728 BP.

XX AC ADO06496;

XX DT 29-JUL-2004 (first entry)

XX DE A thaliana cytokinin oxidase AtCKX1 coding sequence.

XX KW bioremediation; AtCKX1; cytokinin availability; contaminant;  
 KW metal deficiency; nutrition; ds; gene; cytokinin oxidase.

XX OS Arabidopsis thaliana.

XX FH Key    Location/Qualifiers

FT CDS    1..1728

FT    /\*tag= a

FT    /product= "cytokinin oxidase AtCKX1"

XX WO2004038027-A1.

XX PD 06-MAY-2004.

XX PF 24-OCT-2003; 2003WO-EP012051.

XX PR 24-OCT-2002; 2002EP-00079481.

XX PA (CROP-) CROPDESIGN NV.

XX PI Van Camp W;

XX DR WPI; 2004-375913/35.

XX DR P-PSDB; ADO06497.

Method for bioremediation, useful for removing contaminants or metals, by decreasing cytokinin availability in plants, and cultivating plant on substrate comprising one or more contaminants.

Claim 8; Page 53; 61pp; English.

The present invention relates to a method for bioremediation, which involves decreasing cytokinin availability in a plant relative to a corresponding wild type plant, and cultivating the plant on a substrate comprising one or more contaminants, or cultivating a plant having lowered availability of cytokinin relative to corresponding wild type plants, on a substrate, which is to be treated. The method is useful for bioremediation, for concentrating contaminants in a plant, where the plant has a higher concentration of contaminants compared to a corresponding wild type plant. The metal contaminants include aluminum, americium, antimony, arsenic, barium, beryllium, bismuth, cadmium, caesium, cerium, chromium, copper, gallium, germanium, gold, indium, iridium, iron, lead, manganese, mercury, molybdenum, neptunium, osmium, palladium, platinum, plutonium, radium, rhodium, rubidium, ruthenium, scandium, selenium, silver, strontium, technetium, tellurium, thallium, tin, tungsten, uranium, vanadium or yttrium, preferably cadmium. The plant obtained by the method is useful in bioremediation. Transgenic plants with a lower availability of cytokinin are useful in the manufacture of a medicament for treatment of disorders arising from metal deficiencies, and as a medicament for improving animal or human nutrition. The present sequence is a coding sequence shown in the exemplification of the invention.

Sequence 1728 BP; 544 A; 390 C; 332 G; 462 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 601    Length: 1728  
 Score: 35.00    Matches: 7

Percent Similarity: 100.00%      Conservative: 0  
Best Local Similarity: 100.00%      Mismatches: 0  
Query Match: 100.00%      Indels: 0  
DB: 12      Gaps: 0

US-10-014-101B-39 (1-7) x ADO6496 (1-1728)

QY 1 ValGlyGlyThrLeuSerAsn 7  
ID ADH61281  
AC ADH61281;  
DB 574 GTTGGAGGTACACTACTAAT 594

## RESULT 39

ADH61281  
ID ADH61281 standard; DNA; 1857 BP.

XX AC ADH61281;

XX DT 25-MAR-2004 (first entry)

XX DE Barley cytokinin oxidase (CKX1) DNA #1.

XX KW Male-sterility; cytokinin oxidase; transgenic plant; seed production;  
XX KW CKX1; barley; ds; gene.

XX OS Hordeum vulgare.

XX PN US2003163847-A1.

XX PD 28-AUG-2003.

XX PF 20-DEC-2002; 2002US-00326184.

XX PR 20-DEC-2001; 2001US-0343129P.

XX PA (PHAA ) PHARMACIA CORP.

XX PI Huang S, Crossland LD, Cheikh N, Morris RO;

XX WPI; 2003-897983/82.

XX GENBANK; AF540382.

XX PT Producing plants characterized by reversible male-sterility, useful for  
PT maintaining male sterility in plants, by transforming a plant cell with a  
PT nucleic acid construct containing a polynucleotide encoding a cytokinin  
PT oxidase.

XX PS Disclosure; SEQ ID NO 22; 33pp; English.

XX CC The invention relates to a method for producing a plant characterised by  
CC reversible male-sterility which involves transforming a plant cell with a  
CC nucleic acid construct containing a polynucleotide encoding a cytokinin  
CC oxidase. The method is useful for producing reversible male-sterility in  
CC transgenic plants, or for maintaining male sterility in plants. The  
CC method reduces the expense of seed production for existing hybrid plants  
CC such as corn, but also makes it possible to produce hybrid varieties of  
CC traditionally non-hybrid crops. The method is also useful for introducing  
CC economically valuable traits from plants having undesirable production  
CC characteristics into plants having desirable characteristics. The present  
CC sequence is barley cytokinin oxidase (CKX1) DNA. This sequence is used to  
CC illustrate the method of the invention.

XX SQ Sequence 1857 BP; 521 A; 454 C; 437 G; 445 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 651      Length: 1857  
Score: 35.00      Matches: 7  
Percent Similarity: 100.00%      Conservative: 0  
Best Local Similarity: 100.00%      Mismatches: 0  
Query Match: 100.00%      Indels: 0  
DB: 10      Gaps: 0

US-10-014-101B-39 (1-7) x ADH61281 (1-1857)

QY 1 ValGlyGlyThrLeuSerAsn 7  
DB 590 GTTGGGGGCACGTTGTCAAAT 610

## RESULT 40

ADH61280  
ID ADH61280 standard; DNA; 1873 BP.

XX AC ADH61280;

XX DT 25-MAR-2004 (first entry)

XX DE Arabidopsis thaliana cytokinin oxidase (CKX1) DNA #7.

XX KW Male-sterility; cytokinin oxidase; transgenic plant; seed production;  
XX KW CKX1; mouse-ear cress; gene; ds.

XX OS Arabidopsis thaliana.

XX PN US2003163847-A1.

XX PD 28-AUG-2003.

XX PF 20-DEC-2002; 2002US-00326184.

XX PR 20-DEC-2001; 2001US-0343129P.

XX PA (PHAA ) PHARMACIA CORP.

XX PI Huang S, Crossland LD, Cheikh N, Morris RO;

XX WPI; 2003-897983/82.

XX GENBANK; AY091158.

XX PT Producing plants characterized by reversible male-sterility, useful for  
PT maintaining male sterility in plants, by transforming a plant cell with a  
PT nucleic acid construct containing a polynucleotide encoding a cytokinin  
PT oxidase.

XX PS Disclosure; SEQ ID NO 21; 33pp; English.

XX CC The invention relates to a method for producing a plant characterised by  
CC reversible male-sterility which involves transforming a plant cell with a  
CC nucleic acid construct containing a polynucleotide encoding a cytokinin  
CC oxidase. The method is useful for producing reversible male-sterility in  
CC transgenic plants, or for maintaining male sterility in plants. The  
CC method reduces the expense of seed production for existing hybrid plants  
CC such as corn, but also makes it possible to produce hybrid varieties of  
CC traditionally non-hybrid crops. The method is also useful for introducing  
CC economically valuable traits from plants having undesirable production  
CC characteristics into plants having desirable characteristics. The present  
CC sequence is Arabidopsis thaliana cytokinin oxidase (CKX1) DNA. This  
CC sequence is used to illustrate the method of the invention.

XX SQ Sequence 1873 BP; 495 A; 385 C; 482 G; 511 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 657      Length: 1873  
Score: 35.00      Matches: 7  
Percent Similarity: 100.00%      Conservative: 0  
Best Local Similarity: 100.00%      Mismatches: 0  
Query Match: 100.00%      Indels: 0  
DB: 10      Gaps: 0

US-10-014-101B-39 (1-7) x ADH61280 (1-1873)

QY 1 ValGlyGlyThrLeuSerAsn 7

DB 617 GTGGGAGGTACGTTGTCAAAT 637

Search completed: February 18, 2005, 03:26:38  
Job time : 129.02 secs



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 18, 2005, 02:33:38 ; Search time 41.3 Seconds  
(without alignments)  
277.335 Million cell updates/sec

Title: US-10-014-101b-39

Perfect score: 35

Sequence: 1 VGGTUN 7

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Zgapop 6.0 , Zgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=x1h  
-Q=/cgn2\_1/uspto.spool/US10014101/runat.16022005.075846.7970/app.query.fasta\_1.796  
-DB=Issued\_Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORES=ptc -THR MAX=100 -THR MIN=0 -ALIGN=40  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10014101 @CGN 1 124 @runat.16022005.075846.7970 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLCK=100 -LONGLOG  
-DV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTCUS COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	35	100.0	442	US-09-513-999C-2604	Sequence 2604, Ap
2	35	100.0	1605	US-09-124-541-3	Sequence 3, Appli
3	35	100.0	1605	US-09-663-326-3	Sequence 3, Appli
4	35	100.0	6733	US-09-124-541-2	Sequence 2, Appli
5	35	100.0	6733	US-09-663-326-2	Sequence 2, Appli
6	34	97.1	129127	US-09-949-016-13481	Sequence 13481, A
c 7	33	94.3	2828	US-09-949-016-3718	Sequence 3718, Ap
c 8	33	94.3	41754	US-09-949-016-15460	Sequence 15460, A
c 9	32	91.4	548	US-09-270-767-560	Sequence 560, App
c 10	32	91.4	548	US-09-270-767-15842	Sequence 15842, A
11	32	91.4	601	US-09-949-016-163961	Sequence 163961, A
12	32	91.4	601	US-09-949-016-163962	Sequence 163962, A

13	32	91.4	8885	4	US-09-634-238-26	Sequence 26, Appli
14	32	91.4	49931	4	US-09-949-016-13727	Sequence 13727, A
15	32	91.4	49931	4	US-09-949-016-13728	Sequence 13728, A
16	32	91.4	49931	4	US-09-949-016-13729	Sequence 13729, A
17	32	91.4	49971	4	US-09-949-016-16688	Sequence 16688, A
18	32	91.4	94830	4	US-09-949-016-12414	Sequence 12414, A
19	32	91.4	94847	4	US-09-949-016-16336	Sequence 16336, A
20	32	91.4	130298	4	US-09-949-016-16664	Sequence 16664, A
21	32	91.4	4403765	3	US-09-103-840A-2	Sequence 2, Appli
22	32	91.4	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c 23	31	88.6	363	4	US-09-614-221A-253	Sequence 253, App
c 24	31	88.6	473	4	US-09-471-276-723	Sequence 723, App
c 25	31	88.6	506	4	US-09-536-059-5	Sequence 5, Appli
c 26	31	88.6	562	4	US-09-621-976-338	Sequence 338, App
c 27	31	88.6	891	4	US-09-543-681A-1481	Sequence 1481, Ap
28	31	88.6	972	4	US-09-328-352-1982	Sequence 1982, Ap
29	31	88.6	1155	4	US-09-543-681A-2172	Sequence 2172, Ap
30	31	88.6	1273	4	US-09-270-767-14731	Sequence 14731, A
c 31	31	88.6	1831	3	US-09-433-248A-5	Sequence 5, Appli
32	31	88.6	1877	2	US-08-634-924B-1	Sequence 1, Appli
33	31	88.6	1877	4	US-09-395-554-1	Sequence 1, Appli
34	31	88.6	1877	4	US-09-395-554-3	Sequence 3, Appli
c 35	31	88.6	2727	4	US-09-248-796A-638	Sequence 638, App
c 36	31	88.6	14231	3	US-08-961-527-81	Sequence 81, Appli
c 37	31	88.6	19181	4	US-09-949-016-15016	Sequence 15016, A
c 38	31	88.6	26928	4	US-09-544-398B-6	Sequence 6, Appli
c 39	31	88.6	26928	4	US-09-543-771B-6	Sequence 6, Appli
c 40	31	88.6	44342	4	US-09-949-016-12661	Sequence 12661, A
c 41	31	88.6	47787	4	US-09-949-016-11969	Sequence 11969, A
c 42	31	88.6	62804	3	US-09-800-960-3	Sequence 3, Appli
c 43	31	88.6	62804	4	US-10-096-960-3	Sequence 3, Appli
44	31	88.6	76401	4	US-09-949-016-17153	Sequence 17153, A
45	31	88.6	94755	4	US-09-949-016-11839	Sequence 11839, A

# ALIGNMENTS

RESULT 1  
US-09-513-999C-2604/c  
; Sequence 2604, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; FILE REFERENCE: 59 US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 2604  
; LENGTH: 442  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 126..440  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 32  
; OTHER INFORMATION: k=g or t  
US-09-513-999C-2604

Alignment Scores:  
Pred. No.: 26.3 Length: 442  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0

```

DB:                               4          Gaps: 0
US-10-014-101B-39 (1-7) x US-09-513-999C-2604 (1-442)
QY 1 ValGlyGlyThrLeuSerAsn 7
DB 67 GTGGCGGACGCTTCTTAAT 47
RESULT 2
US-09-124-541-3
; Sequence 3, Application US/09124541A
; Patent No. 6229066
; GENERAL INFORMATION:
; APPLICANT: Morris Ph.D., Roy O.
; TITLE OF INVENTION: A CYTOKININ OXIDASE
; FILE REFERENCE: UMO1490
; CURRENT APPLICATION NUMBER: US/09/124,541A
; CURRENT FILING DATE: 1998-07-29
; EARLIER APPLICATION NUMBER: 60/054,268
; EARLIER FILING DATE: 1997-07-30
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1605
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1605)
US-09-124-541-3
Alignment Scores:
Pred. No.: 113          Length: 1605
Score: 35.00           Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3                  Gaps: 0
US-10-014-101B-39 (1-7) x US-09-124-541-3 (1-1605)
QY 1 ValGlyGlyThrLeuSerAsn 7
DB 523 GTGGCGGACGCTGTCCAC 543
RESULT 3
US-09-663-326-3
; Sequence 3, Application US/09663326
; Patent No. 6617497
; GENERAL INFORMATION:
; APPLICANT: Morris Ph.D., Roy O.
; TITLE OF INVENTION: A CYTOKININ OXIDASE
; FILE REFERENCE: UMO1490
; CURRENT APPLICATION NUMBER: US/09/663,326
; CURRENT FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 60/054,268
; PRIOR FILING DATE: 1997-07-30
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1605
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1605)
US-09-663-326-3
Alignment Scores:
Pred. No.: 113          Length: 1605
Score: 35.00           Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

```

```

Query Match: 100.00% Indels: 0
DB: 4              Gaps: 0
US-10-014-101B-39 (1-7) x US-09-663-326-3 (1-1605)
QY 1 ValGlyGlyThrLeuSerAsn 7
DB 523 GTGGCGGACGCTGTCCAC 543
RESULT 4
US-09-124-541-2
; Sequence 2, Application US/09124541A
; Patent No. 6229066
; GENERAL INFORMATION:
; APPLICANT: Morris Ph.D., Roy O.
; TITLE OF INVENTION: A CYTOKININ OXIDASE
; FILE REFERENCE: UMO1490
; CURRENT APPLICATION NUMBER: US/09/124,541A
; CURRENT FILING DATE: 1998-07-29
; EARLIER APPLICATION NUMBER: 60/054,268
; EARLIER FILING DATE: 1997-07-30
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 6733
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)..(6733)
; OTHER INFORMATION: genomic sequence for a cytokinin oxidase from Zea
; OTHER INFORMATION: mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1497)..(2111)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2524)..(3216)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3311)..(3607)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (5697)
US-09-124-541-2
Alignment Scores:
Pred. No.: 571          Length: 6733
Score: 35.00           Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3                  Gaps: 0
US-10-014-101B-39 (1-7) x US-09-124-541-2 (1-6733)
QY 1 ValGlyGlyThrLeuSerAsn 7
DB 2019 GTGGCGGACGCTGTCCAC 2039
RESULT 5
US-09-663-326-2
; Sequence 2, Application US/09663326
; Patent No. 6617497
; GENERAL INFORMATION:
; APPLICANT: Morris Ph.D., Roy O.
; TITLE OF INVENTION: A CYTOKININ OXIDASE
; FILE REFERENCE: UMO1490
; CURRENT APPLICATION NUMBER: US/09/663,326
; CURRENT FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 60/054,268
; PRIOR FILING DATE: 1997-07-30
; NUMBER OF SEQ ID NOS: 20

```

```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 6733
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURES:
; NAME/KEY: gene
; LOCATION: (1)..(6733)
; OTHER INFORMATION: genomic sequence for a cytokinin oxidase from Zea
; OTHER INFORMATION: mays
; NAME/KEY: CDS
; LOCATION: (1497)..(2111)
; NAME/KEY: CDS
; LOCATION: (2524)..(3216)
; NAME/KEY: CDS
; LOCATION: (3311)..(3607)
; NAME/KEY: unsure
; LOCATION: (5697)
US-09-663-326-2

Alignment Scores:
Pred. No.: 571 Length: 6733
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-014-101B-39 (1-7) x US-09-663-326-2 (1-6733)

QY 1 ValGlyGlyThrLeuSerAsn 7
Db 2019 GTCGGCGGCACGCTGTCCAAC 2039

RESULT 6
US-09-949-016-13481
; Sequence 13481, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13481
; LENGTH: 129127
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13481

Alignment Scores:
Pred. No.: 2,73e+04 Length: 129127
Score: 34.00 Matches: 6
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 97.14% Indels: 0
DB: 4 Gaps: 0

US-10-014-101B-39 (1-7) x US-09-949-016-13481 (1-129127)

QY 1 ValGlyGlyThrLeuSerAsn 7
Db 104757 ATAGAGGGACACTTAGCAAT 104777

RESULT 7
US-09-949-016-3718/c
; Sequence 3718, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3718
; LENGTH: 2828
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3718

Alignment Scores:
Pred. No.: 626 Length: 2828
Score: 33.00 Matches: 6
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 85.71% Mismatches: 0
Query Match: 94.29% Indels: 0
DB: 4 Gaps: 0

US-10-014-101B-39 (1-7) x US-09-949-016-3718 (1-2828)

QY 1 ValGlyGlyThrLeuSerAsn 7
Db 1446 GTAGGAGGAACAATATCTAAT 1426

RESULT 8
US-09-949-016-15460/c
; Sequence 15460, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15460
; LENGTH: 41754
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15460

Alignment Scores:
Pred. No.: 1,31e+04 Length: 41754
Score: 33.00 Matches: 6
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 85.71% Mismatches: 0
Query Match: 94.29% Indels: 0
DB: 4 Gaps: 0

US-10-014-101B-39 (1-7) x US-09-949-016-15460 (1-41754)
```

```
QY      1 ValGlyGlyThrLeuSerAsn 7
Db      26813 GTAGGAGGAACAATATCTAAT 26793

RESULT 9
US-09-270-767-560/c
; Sequence 560, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 560
; LENGTH: 548
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-560

Alignment Scores:
Pred. No.:      168      Length:      548
Score:          32.00    Matches:      6
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 85.71% Mismatches: 0
Query Match:     91.43% Indels:      0
DB:              4      Gaps:        0

US-10-014-101B-39 (1-7) x US-09-270-767-560 (1-548)

QY      1 ValGlyGlyThrLeuSerAsn 7
Db      25 GTTGGTGGGACGTTGACTAAT 5

RESULT 10
US-09-270-767-15842/c
; Sequence 15842, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15842
; LENGTH: 548
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-15842

Alignment Scores:
Pred. No.:      168      Length:      548
Score:          32.00    Matches:      6
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 85.71% Mismatches: 0
Query Match:     91.43% Indels:      0
DB:              4      Gaps:        0

US-10-014-101B-39 (1-7) x US-09-270-767-15842 (1-548)

QY      1 ValGlyGlyThrLeuSerAsn 7
Db      25 GTTGGTGGGACGTTGACTAAT 5

RESULT 11
US-09-949-016-163961
; Sequence 163961, Application US/09949016
; Patent No. 6812339
```

```
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 163961
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-163961
```

```
Alignment Scores:
Pred. No.:      186      Length:      601
Score:          32.00    Matches:      6
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 85.71% Mismatches: 0
Query Match:     91.43% Indels:      0
DB:              4      Gaps:        0
```

US-10-014-101B-39 (1-7) x US-09-949-016-163961 (1-601)

```
QY      1 ValGlyGlyThrLeuSerAsn 7
Db      516 GTTGGAGGAACCTTTAGCCAAC 536
```

## RESULT 12

```
US-09-949-016-163962
; Sequence 163962, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 163962
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-163962
```

```
Alignment Scores:
Pred. No.:      186      Length:      601
Score:          32.00    Matches:      6
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 85.71% Mismatches: 0
Query Match:     91.43% Indels:      0
DB:              4      Gaps:        0
```

US-10-014-101B-39 (1-7) x US-09-949-016-163962 (1-601)

```
QY      1 ValGlyGlyThrLeuSerAsn 7
Db      397 GTTGGAGGAACCTTTAGCCAAC 417
```

RESULT 13  
 US-09-634-238-26  
 ; Sequence 26, Application US/09634238  
 ; Patent No. 6544772  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Glenn, Matthew  
 ; APPLICANT: Havukkala, Ilkka J.  
 ; APPLICANT: Bloksberg, Leonard, N.  
 ; APPLICANT: Lubbers, Mark W.  
 ; APPLICANT: Dekker, James  
 ; APPLICANT: Christenson, Anna C.  
 ; APPLICANT: Holland, Ross  
 ; APPLICANT: O'Toole, Paul W.  
 ; APPLICANT: Reid, Julian R.  
 ; APPLICANT: Coolbear, Timothy  
 ; TITLE OF INVENTION: Polynucleotides, materials incorporating  
 ; FILE REFERENCE: 11000.1043U1  
 ; CURRENT APPLICATION NUMBER: US/09/634,238  
 ; CURRENT FILING DATE: 2000-08-08  
 ; NUMBER OF SEQ ID NOS: 422  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 26  
 ; LENGTH: 8885  
 ; TYPE: DNA  
 ; ORGANISM: Lactobacillus rhamnosus  
 US-09-634-238-26

Alignment Scores:  
 Pred. No.: 3.9e+03 Length: 8885  
 Score: 32.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 85.71% Mismatches: 0  
 Query Match: 91.43% Indels: 0  
 DB: 4 Gaps: 0

US-10-014-101B-39 (1-7) x US-09-634-238-26 (1-8885)

QY 1 ValGlyGlyThrLeuSerAsn 7  
 |||||  
 DB 3064 GTGGGTGGAAACGCTACGAC 3084

RESULT 14  
 US-09-949-016-13727  
 ; Sequence 13727, Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 ; FILE REFERENCE: CL001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; CURRENT FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241,755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237,768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231,498  
 ; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 13727  
 ; LENGTH: 49931  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)...(49931)  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-09-949-016-13727

Alignment Scores:

Pred. No.: 2.73e+04 Length: 49931  
 Score: 32.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 85.71% Mismatches: 0  
 Query Match: 91.43% Indels: 0  
 DB: 4 Gaps: 0

US-10-014-101B-39 (1-7) x US-09-949-016-13727 (1-49931)

QY 1 ValGlyGlyThrLeuSerAsn 7  
 |||||  
 DB 20294 GTTGGAGGAACGCTAAACAAT 20314

RESULT 15

US-09-949-016-13728  
 ; Sequence 13728, Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 ; FILE REFERENCE: CL001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; CURRENT FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241,755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237,768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231,498  
 ; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 13728  
 ; LENGTH: 49931  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)...(49931)  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-09-949-016-13728

Alignment Scores:

Pred. No.: 2.73e+04 Length: 49931  
 Score: 32.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 85.71% Mismatches: 0  
 Query Match: 91.43% Indels: 0  
 DB: 4 Gaps: 0

US-10-014-101B-39 (1-7) x US-09-949-016-13728 (1-49931)

QY 1 ValGlyGlyThrLeuSerAsn 7  
 |||||  
 DB 20294 GTTGGAGGAACGCTAAACAAT 20314

RESULT 16  
 US-09-949-016-13729  
 ; Sequence 13729, Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 ; FILE REFERENCE: CL001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; CURRENT FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241,755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237,768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231,498  
 ; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 13727  
 ; LENGTH: 49931  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)...(49931)  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-09-949-016-13727

Alignment Scores:

Pred. No.: 2.73e+04 Length: 49931  
 Score: 32.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 85.71% Mismatches: 0  
 Query Match: 91.43% Indels: 0  
 DB: 4 Gaps: 0

US-10-014-101B-39 (1-7) x US-09-949-016-13728 (1-49931)

QY 1 ValGlyGlyThrLeuSerAsn 7  
 |||||  
 DB 20294 GTTGGAGGAACGCTAAACAAT 20314

RESULT 16

US-09-949-016-13729  
 ; Sequence 13729, Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 ; FILE REFERENCE: CL001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; CURRENT FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241,755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237,768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231,498  
 ; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 13727  
 ; LENGTH: 49931  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)...(49931)  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-09-949-016-13727

Alignment Scores:

Pred. No.: 2.73e+04 Length: 49931  
 Score: 32.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 85.71% Mismatches: 0  
 Query Match: 91.43% Indels: 0  
 DB: 4 Gaps: 0

US-10-014-101B-39 (1-7) x US-09-949-016-13728 (1-49931)

QY 1 ValGlyGlyThrLeuSerAsn 7  
 |||||  
 DB 20294 GTTGGAGGAACGCTAAACAAT 20314

```

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13729
; LENGTH: 49931
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(49931)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13729

Alignment Scores:
Pred. No.: 2,73e+04 49931
Score: 32.00 6
Percent Similarity: 100.00% 1
Best Local Similarity: 85.71% 0
Query Match: 91.43% 0
DB: 4 0
Gaps: 0

US-10-014-101B-39 (1-7) x US-09-949-016-13729 (1-49931)

QY 1 ValGlyGlyThrLeuSerAsn 7
Db 20294 GTTGGAGGACGCTAACAAAT 20314

RESULT 17
US-09-949-016-16688
; Sequence 16688, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16688
; LENGTH: 49971
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(49971)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16688

Alignment Scores:
Pred. No.: 2,73e+04 49971
Score: 32.00 6
Percent Similarity: 100.00% 1
Best Local Similarity: 85.71% 0
Query Match: 91.43% 0
DB: 4 0
Gaps: 0

US-10-014-101B-39 (1-7) x US-09-949-016-16688 (1-49971)

QY 1 ValGlyGlyThrLeuSerAsn 7
Db 32992 CTGGGAGGCACACTTAGTAAC 33012

RESULT 18
US-09-949-016-12414
; Sequence 12414, Application US/09949016
; Patent No. 6812339

```

```

; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12414
; LENGTH: 94830
; TYPE: DNA
; ORGANISM: Human
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12414

Alignment Scores:
Pred. No.: 5,6e+04 94830
Score: 32.00 6
Percent Similarity: 100.00% 1
Best Local Similarity: 85.71% 0
Query Match: 91.43% 0
DB: 4 0
Gaps: 0

US-10-014-101B-39 (1-7) x US-09-949-016-12414 (1-94830)

QY 1 ValGlyGlyThrLeuSerAsn 7
Db 39706 GTTGGAGGACTTAGCCCAAC 39726

RESULT 19
US-09-949-016-16336
; Sequence 16336, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16336
; LENGTH: 94847
; TYPE: DNA
; ORGANISM: Human
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16336

Alignment Scores:
Pred. No.: 5,6e+04 94847
Score: 32.00 6
Percent Similarity: 100.00% 1
Best Local Similarity: 85.71% 0
Query Match: 91.43% 0
DB: 4 0
Gaps: 0

US-10-014-101B-39 (1-7) x US-09-949-016-16336 (1-94847)

QY 1 ValGlyGlyThrLeuSerAsn 7
Db 39706 GTTGGAGGACTTAGCCCAAC 39726

```

## RESULT 20

US-09-949-016-16664  
; Sequence 16664, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16664  
; LENGTH: 130298  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-16664

Alignment Scores:  
Pred. No.: 7.97e+04 Length: 130298  
Score: 32.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 85.71% Mismatches: 0  
Query Match: 91.43% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-39 (1-7) x US-09-949-016-16664 (1-130298)

QY 1 ValGlyGlyThrLeuSerAsn 7

Db 41546 GTGGGGGGAACCTCAACAT 41566

## RESULT 21

US-09-103-840A-2  
; Sequence 2, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; TUBERCULOSIS  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 4403765  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; FEATURE:  
; OTHER INFORMATION: CDC 1551  
; OTHER INFORMATION: "n" bases at various positions throughout the sequence  
; OTHER INFORMATION: represent a, t, c or g  
US-09-103-840A-2

Alignment Scores:  
Pred. No.: 2e+06 Length: 4403765  
Score: 32.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 85.71% Mismatches: 0  
Query Match: 91.43% Indels: 0  
DB: 3 Gaps: 0

US-10-014-101B-39 (1-7) x US-09-103-840A-2 (1-4403765)

QY 1 ValGlyGlyThrLeuSerAsn 7

Db 2484758 GTGGGGGCAAGGTCACTAAC 2484778

## RESULT 22

US-09-103-840A-1  
; Sequence 1, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; TUBERCULOSIS  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 4411529  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; OTHER INFORMATION: H37Rv  
US-09-103-840A-1

Alignment Scores:  
Pred. No.: 2e+06 Length: 4411529  
Score: 32.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 85.71% Mismatches: 0  
Query Match: 91.43% Indels: 0  
DB: 3 Gaps: 0

US-10-014-101B-39 (1-7) x US-09-103-840A-1 (1-4411529)

QY 1 ValGlyGlyThrLeuSerAsn 7

Db 2487514 GTGGGGGCAAGGTCACTAAC 2487534

## RESULT 23

US-09-614-221A-253/c  
; Sequence 253, Application US/09614221A  
; Patent No. 6723837  
; GENERAL INFORMATION:  
; APPLICANT: Karunanandaa, Balasulojini  
; APPLICANT: Yu, Jaehyuk  
; APPLICANT: Kishore, Ganesh M.  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED  
; WITH STEROL SYNTHESIS AND METABOLISM  
; FILE REFERENCE: 16516.075  
; CURRENT APPLICATION NUMBER: US/09/614,221A  
; CURRENT FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/142,981  
; PRIOR FILING DATE: 1999-07-12  
; NUMBER OF SEQ ID NOS: 626  
; SEQ ID NO 253  
; LENGTH: 363  
; TYPE: DNA  
; ORGANISM: Saccharomyces cerevisiae  
US-09-614-221A-253

Alignment Scores:  
Pred. No.: 180 Length: 363  
Score: 31.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 88.57% Indels: 0  
DB: 4 Gaps: 0

```
US-10-014-101B-39 (1-7) x US-09-614-221A-253 (1-363)
QY 2 GlyGlyThrLeuSerAsn 7
Db 199 GGTGGCACCCTGTCCAT 182

RESULT 24
US-09-471-276-723/c
; Sequence 723, Application US/09471276
; Patent No. 6822072
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: GENSET.025CPI
; CURRENT APPLICATION NUMBER: US/09/471,276
; CURRENT FILING DATE: 1999-12-21
; EARLIER APPLICATION NUMBER: 09/057,719
; EARLIER FILING DATE: 1998-04-09
; EARLIER APPLICATION NUMBER: 09/069,047
; EARLIER FILING DATE: 1998-04-28
; EARLIER APPLICATION NUMBER: PCI/IB99/00712
; EARLIER FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 1622
; SOFTWARE: Patent.pm
; SEQ ID NO 723
; LENGTH: 473
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: 63..473
; NAME/KEY: sig_peptide
; LOCATION: 63..212
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 3.5
; OTHER INFORMATION: seq ITYGVFLCICSG/SH
US-09-471-276-723

Alignment Scores:
Pred. No.: 243 Length: 473
Score: 31.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 88.57% Indels: 0
DB: 4 Gaps: 0

US-10-014-101B-39 (1-7) x US-09-471-276-723 (1-473)
QY 2 GlyGlyThrLeuSerAsn 7
Db 219 GGTGGGACCCTGAGCAAT 202

RESULT 25
US-09-536-059-5/c
; Sequence 5, Application US/09536059
; Patent No. 6544737
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Cohen-Akenine, Annick
; TITLE OF INVENTION: GENOMIC SEQUENCE OF THE purH GENE AND purH-RELATED BIALLLELIC
; FILE REFERENCE: GENSET.058AUS
; CURRENT APPLICATION NUMBER: US/09/536,059
; CURRENT FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: US 60/125,961
; PRIOR FILING DATE: 1999-03-24
; NUMBER OF SEQ ID NOS: 24
```

```
; SOFTWARE: Patent.pm
; SEQ ID NO 5
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 72
; OTHER INFORMATION: 99-22580-72 : polymorphic base A or T
; NAME/KEY: misc_binding
; LOCATION: 60..84
; OTHER INFORMATION: 99-22580-72.probe
; NAME/KEY: primer_bind
; LOCATION: 53..71
; OTHER INFORMATION: 99-22580-72.mis
; NAME/KEY: primer_bind
; LOCATION: 73..91
; OTHER INFORMATION: 99-22580-72.mis complement
; NAME/KEY: primer_bind
; LOCATION: 1..18
; OTHER INFORMATION: 99-22580.pu
; NAME/KEY: primer_bind
; LOCATION: 488..506
; OTHER INFORMATION: 99-22580.rp complement
US-09-536-059-5

Alignment Scores:
Pred. No.: 262 Length: 506
Score: 31.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 88.57% Indels: 0
DB: 4 Gaps: 0

US-10-014-101B-39 (1-7) x US-09-536-059-5 (1-506)
QY 2 GlyGlyThrLeuSerAsn 7
Db 332 GGAGGTACTCTCTCTAAC 315

RESULT 26
US-09-621-976-338/c
; Sequence 338, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 338
; LENGTH: 562
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 292..561
; NAME/KEY: sig_peptide
; LOCATION: 292..489
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 3.79999995231628
; OTHER INFORMATION: seq LMLDSCVVPLSP/FP
US-09-621-976-338

Alignment Scores:
Pred. No.: 295 Length: 562
Score: 31.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
```

Query Match: 88.57% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-39 (1-7) x US-09-621-976-338 (1-562)

QY 2 GlyGlyThrLeuSerAsn 7

DB 214 GGTGGGACCTGAGCAAT 197

## RESULT 27

US-09-543-681A-1481/C

; Sequence 1481, Application US/09543681A

; Patent No. 6605709

; GENERAL INFORMATION:

; APPLICANT: GARY BRETON

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS

; FILE REFERENCE: 2709.1002-001

; CURRENT APPLICATION NUMBER: US/09/543.681A

; CURRENT FILING DATE: 2000-04-05

; PRIOR APPLICATION NUMBER: US 60/128,706

; PRIOR FILING DATE: 1999-04-09

; NUMBER OF SEQ ID NOS: 8344

; SEQ ID NO 1481

; LENGTH: 891

; TYPE: DNA

; ORGANISM: Proteus mirabilis

US-09-543-681A-1481

## Alignment Scores:

Pred. No.: 496 Length: 891  
Score: 31.00 Matches: 6  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 88.57%  
DB: 4 Indels: 0  
Gaps: 0

US-10-014-101B-39 (1-7) x US-09-543-681A-1481 (1-891)

QY 2 GlyGlyThrLeuSerAsn 7

DB 743 GCGGACATTGAGTAAC 726

## RESULT 28

US-09-328-352-1982

; Sequence 1982, Application US/09328352

; Patent No. 6562958

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

; FILE REFERENCE: GTC99-03PA

; CURRENT APPLICATION NUMBER: US/09/328.352

; CURRENT FILING DATE: 1999-06-04

; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO 1982

; LENGTH: 972

; TYPE: DNA

; ORGANISM: Acinetobacter baumannii

US-09-328-352-1982

## Alignment Scores:

Pred. No.: 548 Length: 972  
Score: 31.00 Matches: 6  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 88.57%  
DB: 4 Indels: 0  
Gaps: 0

US-10-014-101B-39 (1-7) x US-09-328-352-1982 (1-972)

QY 2 GlyGlyThrLeuSerAsn 7

DB 743 GCGGACATTGAGTAAC 726

DB 613 GGTGGACATTATCGAAT 630

## RESULT 29

US-09-543-681A-2172

; Sequence 2172, Application US/09543681A

; Patent No. 6605709

; GENERAL INFORMATION:

; APPLICANT: GARY BRETON

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS

; FILE REFERENCE: 2709.1002-001

; CURRENT APPLICATION NUMBER: US/09/543.681A

; CURRENT FILING DATE: 2000-04-05

; PRIOR APPLICATION NUMBER: US 60/128,706

; PRIOR FILING DATE: 1999-04-09

; NUMBER OF SEQ ID NOS: 8344

; SEQ ID NO 2172

; LENGTH: 1155

; TYPE: DNA

; ORGANISM: Proteus mirabilis

US-09-543-681A-2172

## Alignment Scores:

Pred. No.: 665 Length: 1155  
Score: 31.00 Matches: 6  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 88.57%  
DB: 4 Indels: 0  
Gaps: 0

US-10-014-101B-39 (1-7) x US-09-543-681A-2172 (1-1155)

QY 2 GlyGlyThrLeuSerAsn 7

DB 988 GGTGGCACATTGAGTAAT 1005

## RESULT 30

US-09-270-767-14731

; Sequence 14731, Application US/09270767

; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

; FILE REFERENCE: File Reference: 7326-094

; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 14731

; LENGTH: 1273

; TYPE: DNA

; ORGANISM: Drosophila melanogaster

US-09-270-767-14731

## Alignment Scores:

Pred. No.: 743 Length: 1273  
Score: 31.00 Matches: 6  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 88.57%  
DB: 4 Indels: 0  
Gaps: 0

US-10-014-101B-39 (1-7) x US-09-270-767-14731 (1-1273)

QY 2 GlyGlyThrLeuSerAsn 7

DB 787 GGAGGTACTCTTCAAT 804

## RESULT 31

US-09-433-248A-5/C

; Sequence 5, Application US/09433248A

; Patent No. 6355462

; GENERAL INFORMATION:

```
; APPLICANT: Falco, S. Carl
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Han, Feng
; APPLICANT: Rafalski, J. Antoni
; TITLE OF INVENTION: Disease Resistance Factors
; FILE REFERENCE: B81252 US NA
; CURRENT APPLICATION NUMBER: US/09/433,248A
; CURRENT FILING DATE: 1999-11-04
; PRIOR APPLICATION NUMBER: 60/107,242
; PRIOR FILING DATE: 1998-11-05
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1831
; TYPE: DNA
; ORGANISM: Glycine max
US-09-433-248A-5

Alignment Scores:
Pred. No.: 1.12e+03 Length: 1831
Score: 31.00 Matches: 5
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 71.43% Mismatches: 0
Query Match: 88.57% Indels: 0
DB: 3 Gaps: 0

US-10-014-101B-39 (1-7) x US-09-433-248A-5 (1-1831)
QY 1 ValGlyGlyThrLeuSerAsn 7
Db 759 ATAGTGGAAGCTGTGACCAAC 739

RESULT 32
US-08-634-924B-1
; Sequence 1, Application US/08634924B
; Patent No. 5834419
; GENERAL INFORMATION:
; APPLICANT: MCFADDEN, GRANT
; APPLICANT: LUCAS, ALEXANDRA
; TITLE OF INVENTION: CHEMOKINE BINDING PROTEIN AND METHODS OF
; FILE REFERENCE: USE THEREFOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Fish & Richardson
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/634,924B
; FILING DATE: 19-APR-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/424,850
; FILING DATE: 19-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Halle, Ph.D., Lisa A.,
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: PD-3675
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1877 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: CBP
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 455...1243
US-08-634-924B-1

Alignment Scores:
Pred. No.: 1.15e+03 Length: 1877
Score: 31.00 Matches: 5
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 71.43% Mismatches: 0
Query Match: 88.57% Indels: 0
DB: 4 Gaps: 0

US-10-014-101B-39 (1-7) x US-09-395-554-1 (1-1877)
QY 1 ValGlyGlyThrLeuSerAsn 7
Db 1180 ATTGGAGGAAGCTGTATCTAAC 1200

RESULT 33
US-09-395-554-1
; Sequence 1, Application US/09395554
; Patent No. 6495515
; GENERAL INFORMATION:
; APPLICANT: MCFADDEN, GRANT
; APPLICANT: LUCAS, ALEXANDRA
; TITLE OF INVENTION: CHEMOKINE BINDING PROTEIN AND METHODS OF
; FILE REFERENCE: 50082/002004
; CURRENT APPLICATION NUMBER: US/09/395,554
; CURRENT FILING DATE: 1999-09-14
; EARLIER APPLICATION NUMBER: 09/181,122
; EARLIER FILING DATE: 1998-10-28
; EARLIER APPLICATION NUMBER: 08/634,924
; EARLIER FILING DATE: 1996-04-19
; EARLIER APPLICATION NUMBER: 08/424,850
; EARLIER FILING DATE: 1995-04-19
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1877
; TYPE: DNA
; ORGANISM: Myxoma virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (455)...(1243)
US-09-395-554-1

Alignment Scores:
Pred. No.: 1.15e+03 Length: 1877
Score: 31.00 Matches: 5
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 71.43% Mismatches: 0
Query Match: 88.57% Indels: 0
DB: 4 Gaps: 0

US-10-014-101B-39 (1-7) x US-09-395-554-1 (1-1877)
QY 1 ValGlyGlyThrLeuSerAsn 7
Db 1180 ATTGGAGGAAGCTGTATCTAAC 1200

RESULT 34
US-09-395-554-3
; Sequence 3, Application US/09395554
; Patent No. 6495515
; GENERAL INFORMATION:
; APPLICANT: MCFADDEN, GRANT
; APPLICANT: LUCAS, ALEXANDRA
```

; TITLE OF INVENTION: CHEMOKINE BINDING PROTEIN AND METHODS OF  
; FILE REFERENCE: 50082/002004  
; CURRENT APPLICATION NUMBER: US/09/395,554  
; CURRENT FILING DATE: 1999-09-14  
; EARLIER APPLICATION NUMBER: 09/181,122  
; EARLIER FILING DATE: 1998-10-28  
; EARLIER APPLICATION NUMBER: 08/634,924  
; EARLIER FILING DATE: 1996-04-19  
; EARLIER APPLICATION NUMBER: 08/424,850  
; EARLIER FILING DATE: 1995-04-19  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 1877  
; TYPE: DNA  
; ORGANISM: Myxoma virus  
US-09-395-554-3

Alignment Scores:  
Pred. No.: 1.15e+03 Length: 1877  
Score: 31.00 Matches: 5  
Percent Similarity: 100.00% Conservative: 2  
Best Local Similarity: 71.43% Mismatches: 0  
Query Match: 88.57% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-39 (1-7) x US-09-395-554-3 (1-1877)

Qy 1 ValGlyGlyThrLeuSerAsn 7  
Db 1180 ATTGGAGGAACTGTATCTAAC 1200  
:::|||||:|||||:|||||

RESULT 35

US-09-248-796A-638/c  
; Sequence 638, Application US/09248796A  
; Patent No. 6747137  
; ORGANISM: Candida albicans

; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstein et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 638  
; LENGTH: 2727  
; TYPE: DNA  
; ORGANISM: Candida albicans

Alignment Scores:  
Pred. No.: 1.76e+03 Length: 2727  
Score: 31.00 Matches: 5  
Percent Similarity: 100.00% Conservative: 2  
Best Local Similarity: 71.43% Mismatches: 0  
Query Match: 88.57% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-39 (1-7) x US-09-248-796A-638 (1-2727)

Qy 1 ValGlyGlyThrLeuSerAsn 7  
Db 2529 ATTGGAGGACGGTAAGTAAAC 2509  
:::|||||:|||||:|||||

RESULT 36

US-08-961-527-81/c  
; Sequence 81, Application US/08961527  
; Patent No. 6420135

; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 391  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/961,527  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB340P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 81:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 14231 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-961-527-81

Alignment Scores:  
Pred. No.: 1.14e+04 Length: 14231  
Score: 31.00 Matches: 5  
Percent Similarity: 100.00% Conservative: 2  
Best Local Similarity: 71.43% Mismatches: 0  
Query Match: 88.57% Indels: 0  
DB: 3 Gaps: 0

US-10-014-101B-39 (1-7) x US-08-961-527-81 (1-14231)

Qy 1 ValGlyGlyThrLeuSerAsn 7  
Db 5528 ATGTGTGAACCTTAGCAAC 5508  
:::|||||:|||||:|||||

RESULT 37

US-09-949-016-15016  
; Sequence 15016, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15016  
; LENGTH: 19181

**TITLE OF INVENTION:** High bone mass gene of 11q13.3

16570 GGTGGACCTGAGCAAT 16553

Search completed: February 18, 2005, 07:16:11  
Job time : 586.55 secs

---

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 18, 2005, 05:26:14 ; Search time 329.98 Seconds  
(without alignments)  
125.382 Million cell updates/sec

Title: US-10-014-101B-39

Perfect score: 35

Sequence: 1 VGGTLSN 7

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 5384158 seqs, 2955248155 residues

Total number of hits satisfying chosen parameters: 10768316

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US10014101/runat\_16022005\_075848\_8061/app.query.fasta\_1.796  
-DB=Published Applications NA -QWMT=fastap -SUPFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=40 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10014101 @CGN 1 659 @runat 16022005\_075848\_8061  
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq:  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:  
18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq:  
19: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:  
20: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:  
21: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:  
22: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	35	100.0	470	18	US-10-425-115-168063	Sequence 168063,
2	35	100.0	566	16	US-10-326-184-14	Sequence 14, Appl
3	35	100.0	579	18	US-10-021-323-14076	Sequence 14076, A
4	35	100.0	641	18	US-10-021-323-8905	Sequence 8905, Ap
5	35	100.0	718	17	US-10-424-599-69520	Sequence 69520, A
6	35	100.0	840	18	US-10-425-115-35579	Sequence 35579, A
7	35	100.0	955	18	US-10-425-115-35581	Sequence 35581, A
8	35	100.0	1142	17	US-10-425-114-8251	Sequence 8251, Ap
9	35	100.0	1318	17	US-10-424-599-9802	Sequence 9802, Ap
10	35	100.0	1318	17	US-10-424-599-12845	Sequence 12845, A
11	35	100.0	1384	17	US-10-425-114-8799	Sequence 8799, Ap
12	35	100.0	1506	14	US-10-014-101-26	Sequence 26, Appl
13	35	100.0	1506	17	US-10-311-453-26	Sequence 26, Appl
14	35	100.0	1515	14	US-10-014-101-30	Sequence 30, Appl
15	35	100.0	1515	17	US-10-311-453-30	Sequence 30, Appl
16	35	100.0	1560	18	US-10-817-483-26	Sequence 26, Appl
17	35	100.0	1566	18	US-10-817-483-30	Sequence 30, Appl
18	35	100.0	1572	14	US-10-014-101-27	Sequence 27, Appl
19	35	100.0	1572	16	US-10-326-184-9	Sequence 9, Appl
20	35	100.0	1572	17	US-10-311-453-27	Sequence 27, Appl
21	35	100.0	1575	9	US-09-938-842A-2089	Sequence 2089, Ap
22	35	100.0	1575	11	US-09-938-842A-2089	Sequence 2089, Ap
23	35	100.0	1575	14	US-10-014-101-28	Sequence 28, Appl
24	35	100.0	1575	16	US-10-326-184-10	Sequence 10, Appl
25	35	100.0	1575	16	US-10-326-184-11	Sequence 11, Appl
26	35	100.0	1575	17	US-10-311-453-28	Sequence 28, Appl
27	35	100.0	1599	18	US-10-437-963-88564	Sequence 88564, A
28	35	100.0	1605	18	US-10-437-963-88170	Sequence 88170, A
29	35	100.0	1605	18	US-10-821-711-4	Sequence 4, Appl
30	35	100.0	1605	18	US-10-821-711-5	Sequence 5, Appl
31	35	100.0	1611	14	US-10-014-101-29	Sequence 29, Appl
32	35	100.0	1611	17	US-10-311-453-29	Sequence 29, Appl
33	35	100.0	1617	18	US-10-817-483-28	Sequence 28, Appl
34	35	100.0	1620	14	US-10-014-101-34	Sequence 34, Appl
35	35	100.0	1620	16	US-10-311-453-34	Sequence 34, Appl
36	35	100.0	1623	16	US-10-326-184-12	Sequence 12, Appl
37	35	100.0	1629	18	US-10-817-483-32	Sequence 32, Appl
38	35	100.0	1655	16	US-10-326-184-20	Sequence 20, Appl
39	35	100.0	1670	18	US-10-437-963-70571	Sequence 70571, A
40	35	100.0	1673	18	US-10-437-963-83827	Sequence 83827, A
41	35	100.0	1677	18	US-10-437-963-43981	Sequence 43981, A
42	35	100.0	1718	18	US-10-437-963-86803	Sequence 86803, A
43	35	100.0	1719	17	US-10-369-493-27883	Sequence 27883, A
44	35	100.0	1728	14	US-10-014-101-25	Sequence 25, Appl
45	35	100.0	1728	17	US-10-311-453-25	Sequence 25, Appl

#### ALIGNMENTS

RESULT 1  
US-10-425-115-168063  
; Sequence 168063, Application US/10425115  
; Publication No. US2004021472A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 168063  
; LENGTH: 470  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:

; NAME/KEY: unsure  
; LOCATION: (1)..(470)  
; OTHER INFORMATION: unsure at all n locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MET4577\_84854C.1  
US-10-425-115-168063

Alignment Scores:  
Pred. No.: 156 Length: 470  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 18 Gaps: 0

US-10-014-101B-39 (1-7) x US-10-425-115-168063 (1-470)

Qy 1 ValGlyGlyThrLeuSerAsn 7  
Db 295 GTGGCGGACGCTCTCCAAT 315

## RESULT 2

US-10-326-184-14  
; Sequence 14, Application US/10326184  
; Publication No. US20030163847A1  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Company  
; TITLE OF INVENTION: REVERSIBLE MALE STERILITY IN TRANSGENIC PLANTS BY EXPRESSION OF  
; FILE REFERENCE: MTC6781.1  
; CURRENT APPLICATION NUMBER: US/10/326,184  
; CURRENT FILING DATE: 2002-12-20  
; PRIOR APPLICATION NUMBER: US 60/343,129  
; PRIOR FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 14  
; LENGTH: 566  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-10-326-184-14

Alignment Scores:  
Pred. No.: 188 Length: 566  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 16 Gaps: 0

US-10-014-101B-39 (1-7) x US-10-326-184-14 (1-566)

Qy 1 ValGlyGlyThrLeuSerAsn 7  
Db 133 GTGGAGGACACTTCTCCAAT 153

## RESULT 3

US-10-021-323-14076  
; Sequence 14076, Application US/10021323  
; Publication No. US20040123340A1  
; GENERAL INFORMATION:  
; APPLICANT: Deikman, Jill  
; APPLICANT: Feng, Paul C.C.  
; APPLICANT: Fincher, Karen L.  
; APPLICANT: Ziegler, Todd E.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(52274)B  
; CURRENT APPLICATION NUMBER: US/10/021,323  
; CURRENT FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: US 60/255, 619  
; PRIOR FILING DATE: 2000-12-14  
; NUMBER OF SEQ ID NOS: 17880

; SEQ ID NO 14076  
; LENGTH: 579  
; TYPE: DNA  
; ORGANISM: Gossypium hirsutum  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(579)  
; OTHER INFORMATION: unsure at all n locations  
; OTHER INFORMATION: Clone ID: LIB3829-029-Q6-K6-G4  
US-10-021-323-14076

Alignment Scores:  
Pred. No.: 193 Length: 579  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 18 Gaps: 0

US-10-014-101B-39 (1-7) x US-10-021-323-14076 (1-579)

Qy 1 ValGlyGlyThrLeuSerAsn 7  
Db 209 GTTGTGGAACCTTGCTAAC 229

## RESULT 4

US-10-021-323-8905  
; Sequence 8905, Application US/10021323  
; Publication No. US20040123340A1  
; GENERAL INFORMATION:  
; APPLICANT: Deikman, Jill  
; APPLICANT: Feng, Paul C.C.  
; APPLICANT: Fincher, Karen L.  
; APPLICANT: Ziegler, Todd E.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(52274)B  
; CURRENT APPLICATION NUMBER: US/10/021,323  
; CURRENT FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: US 60/255, 619  
; PRIOR FILING DATE: 2000-12-14  
; NUMBER OF SEQ ID NOS: 17880  
; SEQ ID NO 8905  
; LENGTH: 641  
; TYPE: DNA  
; ORGANISM: Gossypium hirsutum  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3828-021-Q1-K6-C8  
US-10-021-323-8905

Alignment Scores:  
Pred. No.: 213 Length: 641  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 18 Gaps: 0

US-10-014-101B-39 (1-7) x US-10-021-323-8905 (1-641)

Qy 1 ValGlyGlyThrLeuSerAsn 7  
Db 541 GTTGTGGAACCTTGCTAAC 561

## RESULT 5

US-10-424-599-69520  
; Sequence 69520, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei

```
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 69520
; LENGTH: 718
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_33789C.1
US-10-424-599-69520

Alignment Scores:
Pred. No.: 239 Length: 718
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-10-014-101B-39 (1-7) x US-10-424-599-69520 (1-718)

QY 1 ValGlyGlyThrLeuSerAen 7
Db 650 GTTGGTGGCACTCTCTCCAC 670

RESULT 6
US-10-425-115-35579
; Sequence 35579, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 35579
; LENGTH: 840
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(840)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_132446C.1
US-10-425-115-35579

Alignment Scores:
Pred. No.: 279 Length: 840
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-10-014-101B-39 (1-7) x US-10-425-115-35579 (1-840)

QY 1 ValGlyGlyThrLeuSerAen 7
Db 497 GTCGGTGGCACTTGTCTAAT 517

RESULT 7
US-10-425-115-35581
; Sequence 35581, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 35581
; LENGTH: 955
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(955)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_132448C.1
US-10-425-115-35581

Alignment Scores:
Pred. No.: 318 Length: 955
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-10-014-101B-39 (1-7) x US-10-425-115-35581 (1-955)

QY 1 ValGlyGlyThrLeuSerAen 7
Db 680 GTCGGTGGCACTTGTCTAAT 700

RESULT 8
US-10-425-114-8251
; Sequence 8251, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 8251
; LENGTH: 1142
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700742678_FLI
US-10-425-114-8251

Alignment Scores:
Pred. No.: 380 Length: 1142
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-10-014-101B-39 (1-7) x US-10-425-114-8251 (1-1142)

QY 1 ValGlyGlyThrLeuSerAen 7
```

```
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 35581
; LENGTH: 955
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(955)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_132448C.1
US-10-425-115-35581

Alignment Scores:
Pred. No.: 318 Length: 955
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-10-014-101B-39 (1-7) x US-10-425-115-35581 (1-955)

QY 1 ValGlyGlyThrLeuSerAen 7
Db 680 GTCGGTGGCACTTGTCTAAT 700

RESULT 8
US-10-425-114-8251
; Sequence 8251, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 8251
; LENGTH: 1142
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700742678_FLI
US-10-425-114-8251

Alignment Scores:
Pred. No.: 380 Length: 1142
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-10-014-101B-39 (1-7) x US-10-425-114-8251 (1-1142)

QY 1 ValGlyGlyThrLeuSerAen 7
```

```
Db      129  GTGGGGGACGCTCTCCAAT 149
|||||
RESULT 9
US-10-424-599-9802
; Sequence 9802, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 9802
; LENGTH: 1318
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_108862C.1
US-10-424-599-9802

Alignment Scores:
Pred. No.:      438      Length:      1318
Score:          35.00    Matches:      7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:     100.00% Indels: 0
DB:              17      Gaps: 0

US-10-014-101B-39 (1-7) x US-10-424-599-9802 (1-1318)

QY      1  ValGlyGlyThrLeuSerAsn 7
|||||
Db      628  GTGGGGGACGCTCTCCAAT 648
|||||
RESULT 10
US-10-424-599-12845
; Sequence 12845, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 12845
; LENGTH: 1318
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1318)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_111605C.1
US-10-424-599-12845

Alignment Scores:
Pred. No.:      438      Length:      1318
Score:          35.00    Matches:      7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:     100.00% Indels: 0
DB:              17      Gaps: 0

US-10-014-101B-39 (1-7) x US-10-424-599-9802 (1-1318)

QY      1  ValGlyGlyThrLeuSerAsn 7
|||||
Db      628  GTGGGGGACGCTCTCCAAT 648
|||||
RESULT 11
US-10-425-114-8799
; Sequence 8799, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 8799
; LENGTH: 1384
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700793982_FLI
US-10-425-114-8799

Alignment Scores:
Pred. No.:      460      Length:      1384
Score:          35.00    Matches:      7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:     100.00% Indels: 0
DB:              17      Gaps: 0

US-10-014-101B-39 (1-7) x US-10-425-114-8799 (1-1384)

QY      1  ValGlyGlyThrLeuSerAsn 7
|||||
Db      92  GTGGGGGACGCTCTCCAAT 112
|||||
RESULT 12
US-10-014-101-26
; Sequence 26, Application US/10014101
; Publication No. US20030074698A1
; GENERAL INFORMATION:
; APPLICANT: Schmullling, Thomas
; APPLICANT: Werner, Tom B
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and
; FILE REFERENCE: 1195-2
; CURRENT APPLICATION NUMBER: US/10/014,101
; CURRENT FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: PCT/EP01/06833
; PRIOR FILING DATE: 2001-06-16
; PRIOR APPLICATION NUMBER: EP 00870132.8
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/258,415
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: EP 01870053.4
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 1506
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
```

```
US-10-014-101-26
Alignment Scores:
Pred. No.: 501 Length: 1506
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-10-014-101B-39 (1-7) x US-10-014-101-26 (1-1506)
QY 1 ValGlyGlyThrluSerAen 7
Db 466 GTCGAGGAACGTGTGCAAT 486

RESULT 13
US-10-311-453-26
; Sequence 26, Application US/10311453
; Publication No. US20040031073A1
; GENERAL INFORMATION:
; APPLICANT: Schmulling, Thomas
; APPLICANT: Werner, Tom 8
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and
; FILE OF INVENTION: physiology
; FILE REFERENCE: 1226-4
; CURRENT APPLICATION NUMBER: US/10/311,453
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: EP 00870132.8
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/258,415
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: EP 01870053.4
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-311-453-26
Alignment Scores:
Pred. No.: 501 Length: 1506
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-10-014-101B-39 (1-7) x US-10-311-453-26 (1-1506)
QY 1 ValGlyGlyThrluSerAen 7
Db 466 GTCGAGGAACGTGTGCAAT 486

RESULT 14
US-10-014-101-30
; Sequence 30, Application US/10014101
; Publication No. US20030074698A1
; GENERAL INFORMATION:
; APPLICANT: Schmulling, Thomas
; APPLICANT: Werner, Tom 8
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and
; FILE OF INVENTION: physiology
; FILE REFERENCE: 1195-2
; CURRENT APPLICATION NUMBER: US/10/014,101
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: PCT/EP01/06833
; PRIOR FILING DATE: 2001-06-16
; PRIOR APPLICATION NUMBER: EP 00870132.8
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/258,415
```

```

; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: EP 01870053.4
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 1515
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-014-101-30
Alignment Scores:
Pred. No.: 504 Length: 1515
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-10-014-101B-39 (1-7) x US-10-014-101-30 (1-1515)
QY 1 ValGlyGlyThrluSerAen 7
Db 493 GTAGGTGGTACTCTGTCCAAT 513

RESULT 15
US-10-311-453-30
; Sequence 30, Application US/10311453
; Publication No. US20040031073A1
; GENERAL INFORMATION:
; APPLICANT: Schmulling, Thomas
; APPLICANT: Werner, Tom 8
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and
; FILE OF INVENTION: physiology
; FILE REFERENCE: 1226-4
; CURRENT APPLICATION NUMBER: US/10/311,453
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: EP 00870132.8
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/258,415
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: EP 01870053.4
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 1515
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-311-453-30
Alignment Scores:
Pred. No.: 504 Length: 1515
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-10-014-101B-39 (1-7) x US-10-311-453-30 (1-1515)
QY 1 ValGlyGlyThrluSerAen 7
Db 493 GTAGGTGGTACTCTGTCCAAT 513

RESULT 16
US-10-817-483-26
; Sequence 26, Application US/10817483
; Publication No. US20040237147A1
; GENERAL INFORMATION:
; APPLICANT: Habben, Jeffrey E.
; APPLICANT: Zinsmeister, Christopher
; APPLICANT: Tomes, Dwight
```

```

; APPLICANT: Abbott, Shane
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Niu, Xiaomu
; TITLE OF INVENTION: Modulation of Cytokinin Activity in
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 0803R
; CURRENT APPLICATION NUMBER: US/10/817,483
; CURRENT FILING DATE: 2004-04-02
; PRIOR APPLICATION NUMBER: US 60/460,718
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US 09/545,334
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: US 60/129,844
; PRIOR FILING DATE: 1999-04-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 1560
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1566)
US-10-817-483-26

```

```

Alignment Scores:
Pred. No.: 519 Length: 1560
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

```

US-10-014-101B-39 (1-7) x US-10-817-483-26 (1-1560)

```

QY 1 ValGlyGlyThrLeuSerAsn 7
Db 490 GTGGTGGGCGACCTGTCTCAAT 510

```

RESULT 17

```

US-10-817-483-30
; Sequence 30, Application US/10817483
; Publication No. US20040237147A1
; GENERAL INFORMATION:
; APPLICANT: Habben, Jeffrey E.
; APPLICANT: Zinselmeier, Christopher
; APPLICANT: Tones, Dwight
; APPLICANT: Abbott, Shane
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Niu, Xiaomu
; TITLE OF INVENTION: Modulation of Cytokinin Activity in
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 0803R
; CURRENT APPLICATION NUMBER: US/10/817,483
; CURRENT FILING DATE: 2004-04-02
; PRIOR APPLICATION NUMBER: US 60/460,718
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US 09/545,334
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: US 60/129,844
; PRIOR FILING DATE: 1999-04-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 1566
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1566)
US-10-817-483-30

```

Alignment Scores:

```

Pred. No.: 521 Length: 1566
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

```

US-10-014-101B-39 (1-7) x US-10-817-483-30 (1-1566)

```

QY 1 ValGlyGlyThrLeuSerAsn 7
Db 469 GTGGGCGGCGACGCTCTCAAT 489

```

RESULT 18

```

US-10-014-101-27
; Sequence 27, Application US/10014101
; Publication No. US20030074698A1
; GENERAL INFORMATION:
; APPLICANT: Schmulling, Thomas
; APPLICANT: Werner, Tom S
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and
; TITLE OF INVENTION: Physiology
; FILE REFERENCE: 1195-2
; CURRENT APPLICATION NUMBER: US/10/014,101
; CURRENT FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: PCT/EP01/06833
; PRIOR FILING DATE: 2001-06-16
; PRIOR APPLICATION NUMBER: EP 00870132.8
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/258,415
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: EP 01870053.4
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 1572
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-014-101-27

```

Alignment Scores:

```

Pred. No.: 523 Length: 1572
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

```

US-10-014-101B-39 (1-7) x US-10-014-101-27 (1-1572)

```

QY 1 ValGlyGlyThrLeuSerAsn 7
Db 517 GTGGTGGGCGCTTATCAAC 537

```

RESULT 19

```

US-10-326-184-9
; Sequence 9, Application US/10326184
; Publication No. US20030163847A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Company
; TITLE OF INVENTION: REVERSIBLE MALE STERILITY IN TRANSGENIC PLANTS BY EXPRESSION OF CY
; FILE REFERENCE: MTC6781.1
; CURRENT APPLICATION NUMBER: US/10/326,184
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US 60/343,129
; PRIOR FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 1572
; TYPE: DNA

```

; ORGANISM: Arabidopsis thaliana  
US-10-326-184-9

Alignment Scores: Length: 1572  
Pred. No.: 523 Matches: 7  
Score: 35.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 16 Indels: 0  
Gaps: 0

US-10-014-101B-39 (1-7) x US-10-326-184-9 (1-1572)

QY 1 ValGlyGlyThrLeuSerAen 7  
Db 517 GTCGGTGGGACGTTATCAAC 537

## RESULT 20

US-10-311-453-27  
; Sequence 27, Application US/10311453  
; Publication No. US20040031073A1  
; GENERAL INFORMATION:  
; APPLICANT: Schmulling, Thomas  
; APPLICANT: Werner, Tom S  
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and  
; FILE REFERENCE: 1226-4  
; CURRENT APPLICATION NUMBER: US/10/311.453  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: EP 00870132.8  
; PRIOR FILING DATE: 2000-06-16  
; PRIOR APPLICATION NUMBER: US 60/258,415  
; PRIOR FILING DATE: 2000-12-27  
; PRIOR APPLICATION NUMBER: EP 01870053.4  
; PRIOR FILING DATE: 2001-03-16  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn ver. 2.1  
; SEQ ID NO 27  
; LENGTH: 1572  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-10-311-453-27

Alignment Scores: Length: 1572  
Pred. No.: 523 Matches: 7  
Score: 35.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 17 Indels: 0  
Gaps: 0

US-10-014-101B-39 (1-7) x US-10-311-453-27 (1-1572)

QY 1 ValGlyGlyThrLeuSerAen 7  
Db 517 GTCGGTGGGACGTTATCAAC 537

## RESULT 21

US-09-938-842A-2089  
; Sequence 2089, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SCRIP1300-3  
; CURRENT APPLICATION NUMBER: US/09/938.842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24

; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 2089  
; LENGTH: 1575  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-2089

Alignment Scores: Length: 1575  
Pred. No.: 524 Matches: 7  
Score: 35.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 9 Indels: 0  
Gaps: 0

US-10-014-101B-39 (1-7) x US-09-938-842A-2089 (1-1575)

QY 1 ValGlyGlyThrLeuSerAen 7  
Db 535 GTCGGGGGACGTTGCGAAC 555

## RESULT 22

US-09-938-842A-2089  
; Sequence 2089, Application US/09938842A  
; Publication No. US20040009476A9  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SCRIP1300-3  
; CURRENT APPLICATION NUMBER: US/09/938.842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 2089  
; LENGTH: 1575  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-2089

Alignment Scores: Length: 1575  
Pred. No.: 524 Matches: 7  
Score: 35.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 11 Indels: 0  
Gaps: 0

US-10-014-101B-39 (1-7) x US-09-938-842A-2089 (1-1575)

QY 1 ValGlyGlyThrLeuSerAen 7  
Db 535 GTCGGGGGACGTTGCGAAC 555

## RESULT 23

US-10-014-101-28  
; Sequence 28, Application US/10014101  
; Publication No. US20030074698A1  
; GENERAL INFORMATION:  
; APPLICANT: Schmulling, Thomas  
; APPLICANT: Werner, Tom S

```
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and
; FILE REFERENCE: 1195-2
; CURRENT APPLICATION NUMBER: US/10/014,101
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: PCT/EP01/06833
; PRIOR FILING DATE: 2001-06-16
; PRIOR APPLICATION NUMBER: EP 00870132.8
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/258,415
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: EP 01870053.4
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 1575
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-10-014-101-28

Alignment Scores:
Pred. No.: 524 Length: 1575
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-10-014-101B-39 (1-7) x US-10-014-101-28 (1-1575)

QY 1 ValGlyGlyThrLeuSerAsn 7
Db 535 GTGCGGGGACGTTGTCGAAC 555

RESULT 24
US-10-326-184-10
; Sequence 10, Application US/10326184
; Publication No. US20030163847A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Company
; TITLE OF INVENTION: REVERSIBLE MALE STERILITY IN TRANSGENIC PLANTS BY EXPRESSION OF
; FILE REFERENCE: MTC6781.1
; CURRENT APPLICATION NUMBER: US/10/326,184
; CURRENT FILING DATE: 2002-12-20
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 60/343,129
; PRIOR FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 1575
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-10-326-184-10

Alignment Scores:
Pred. No.: 524 Length: 1575
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 16 Gaps: 0

US-10-014-101B-39 (1-7) x US-10-326-184-10 (1-1575)

QY 1 ValGlyGlyThrLeuSerAsn 7
Db 535 GTGCGGGGACGTTGTCGAAC 555

RESULT 25
US-10-326-184-11
; Sequence 11, Application US/10326184
```

```
; Publication No. US20030163847A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Company
; TITLE OF INVENTION: REVERSIBLE MALE STERILITY IN TRANSGENIC PLANTS BY EXPRESSION OF C
; FILE REFERENCE: MTC6781.1
; CURRENT APPLICATION NUMBER: US/10/326,184
; CURRENT FILING DATE: 2002-12-20
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 60/343,129
; PRIOR FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 1575
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-10-326-184-11

Alignment Scores:
Pred. No.: 524 Length: 1575
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 16 Gaps: 0

US-10-014-101B-39 (1-7) x US-10-326-184-11 (1-1575)

QY 1 ValGlyGlyThrLeuSerAsn 7
Db 502 GTGGGAGGTACGTTGTCAAAT 522

RESULT 26
US-10-311-453-28
; Sequence 28, Application US/10311453
; Publication No. US20040031073A1
; GENERAL INFORMATION:
; APPLICANT: Schmilling, Thomas
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and
; FILE REFERENCE: 1226-4
; CURRENT APPLICATION NUMBER: US/10/311,453
; CURRENT FILING DATE: 2002-12-16
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: EP 00870132.8
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/258,415
; PRIOR FILING DATE: 2000-12-27
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 1575
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-10-311-453-28

Alignment Scores:
Pred. No.: 524 Length: 1575
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-10-014-101B-39 (1-7) x US-10-311-453-28 (1-1575)

QY 1 ValGlyGlyThrLeuSerAsn 7
Db 535 GTGCGGGGACGTTGTCGAAC 555

RESULT 27
```

US-10-437-963-88564  
; Sequence 88564, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 88564  
; LENGTH: 1599  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_87402C.1  
US-10-437-963-88564

Alignment Scores:  
Pred. No.: 532 Length: 1599  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 18 Gaps: 0

US-10-014-101B-39 (1-7) x US-10-437-963-88564 (1-1599)

QY 1 ValGlyGlyThrLeuSerAsn 7  
DB 502 GTAGCGCGGACGCTGTCCAAC 522

RESULT 28  
US-10-437-963-38170  
; Sequence 38170, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 38170  
; LENGTH: 1605  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_41832C.1  
US-10-437-963-38170

Alignment Scores:  
Pred. No.: 534 Length: 1605  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0

DB: 18 Gaps: 0

US-10-014-101B-39 (1-7) x US-10-437-963-38170 (1-1605)

QY 1 ValGlyGlyThrLeuSerAsn 7  
DB 517 GTGGCGGCGACCCCTCTCCAAC 537

RESULT 29  
US-10-821-711-4  
; Sequence 4, Application US/10821711  
; Publication No. US20040237142A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbertson, Larry A  
; APPLICANT: Krieger, Elysia K  
; APPLICANT: Ye, Xudong  
; APPLICANT: Zhang, Wanggen  
; TITLE OF INVENTION: DNA CONSTRUCTS AND METHODS TO ENHANCE THE PRODUCTION OF  
; FILE REFERENCE: 38-21(52967)B  
; CURRENT APPLICATION NUMBER: US/10/821,711  
; CURRENT FILING DATE: 2004-04-08  
; PRIOR APPLICATION NUMBER: 60/461,459  
; PRIOR FILING DATE: 2003-04-09  
; NUMBER OF SEQ ID NOS: 16  
; SEQ ID NO 4  
; LENGTH: 1605  
; TYPE: DNA  
; ORGANISM: Zea mays  
US-10-821-711-4

Alignment Scores:  
Pred. No.: 534 Length: 1605  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 18 Gaps: 0

US-10-014-101B-39 (1-7) x US-10-821-711-4 (1-1605)

QY 1 ValGlyGlyThrLeuSerAsn 7  
DB 523 GTCCGCGGCGACGCTGTCCAAC 543

RESULT 30  
US-10-821-711-5  
; Sequence 5, Application US/10821711  
; Publication No. US20040237142A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbertson, Larry A  
; APPLICANT: Krieger, Elysia K  
; APPLICANT: Ye, Xudong  
; APPLICANT: Zhang, Wanggen  
; TITLE OF INVENTION: DNA CONSTRUCTS AND METHODS TO ENHANCE THE PRODUCTION OF  
; FILE REFERENCE: 38-21(52967)B  
; CURRENT APPLICATION NUMBER: US/10/821,711  
; CURRENT FILING DATE: 2004-04-08  
; PRIOR APPLICATION NUMBER: 60/461,459  
; PRIOR FILING DATE: 2003-04-09  
; NUMBER OF SEQ ID NOS: 16  
; SEQ ID NO 5  
; LENGTH: 1605  
; TYPE: DNA  
; ORGANISM: Bacillus subtilis  
US-10-821-711-5

Alignment Scores:  
Pred. No.: 534 Length: 1605  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0

```
Query Match: 100.0% Indels: 0
DB: 18 Gaps: 0
US-10-014-101B-39 (1-7) x US-10-821-711-5 (1-1605)
QY 1 ValGlyGlyThrLeuSerAsn 7
DB 523 GTCGGCGGCACGCTCTCCAA 543
RESULT 31
US-10-014-101-29
; Sequence 29, Application US/10014101
; Publication No. US20030074698A1
; GENERAL INFORMATION:
; APPLICANT: Schmulling, Thomas
; APPLICANT: Werner, Tom s
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and
; FILE REFERENCE: 1195-2
; CURRENT APPLICATION NUMBER: US/10/014,101
; CURRENT FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: PCT/EP01/06833
; PRIOR FILING DATE: 2001-06-16
; PRIOR APPLICATION NUMBER: EP 00870132.8
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/258,415
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: EP 01870053.4
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 1611
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-014-101-29
Alignment Scores:
Pred. No.: 536 Length: 1611
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0
US-10-014-101B-39 (1-7) x US-10-014-101-29 (1-1611)
QY 1 ValGlyGlyThrLeuSerAsn 7
DB 499 GTTGAGGTACACTCTCCAA 519
RESULT 32
US-10-311-453-29
; Sequence 29, Application US/10311453
; Publication No. US20040031073A1
; GENERAL INFORMATION:
; APPLICANT: Schmulling, Thomas
; APPLICANT: Werner, Tom s
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and
; FILE REFERENCE: 1226-4
; CURRENT APPLICATION NUMBER: US/10/311,453
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: EP 00870132.8
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/258,415
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: EP 01870053.4
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 1611
```

```
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-311-453-29
Alignment Scores:
Pred. No.: 536 Length: 1611
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0
US-10-014-101B-39 (1-7) x US-10-311-453-29 (1-1611)
QY 1 ValGlyGlyThrLeuSerAsn 7
DB 499 GTTGAGGTACACTCTCCAA 519
RESULT 33
US-10-817-483-28
; Sequence 28, Application US/10817483
; Publication No. US20040237147A1
; GENERAL INFORMATION:
; APPLICANT: Habben, Jeffrey E.
; APPLICANT: Zinselmeier, Christopher
; APPLICANT: Tomes, Dwight
; APPLICANT: Abbott, Shane
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Niu, Xiaomu
; TITLE OF INVENTION: Modulation of Cytokinin Activity in
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 0803R
; CURRENT APPLICATION NUMBER: US/10/817,483
; CURRENT FILING DATE: 2004-04-02
; PRIOR APPLICATION NUMBER: US 60/460,718
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US 09/545,334
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: US 60/129,844
; PRIOR FILING DATE: 1999-04-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 1617
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1617)
US-10-817-483-28
Alignment Scores:
Pred. No.: 538 Length: 1617
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0
US-10-014-101B-39 (1-7) x US-10-817-483-28 (1-1617)
QY 1 ValGlyGlyThrLeuSerAsn 7
DB 517 GTCGGCGGCACGCTCTCCAA 537
RESULT 34
US-10-014-101-34
; Sequence 34, Application US/10014101
; Publication No. US20030074698A1
; GENERAL INFORMATION:
; APPLICANT: Schmulling, Thomas
; APPLICANT: Werner, Tom s
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and
```

RESULT 35  
US-10-014-101B-39 (1-7) x US-10-014-101-34 (1-1620)  
QY 1 ValGlyGlyThrLeuSerAan 7  
DB 508 GTTGGAGGTACACTCTCCAAT 528  
US-10-311-453-34  
; Sequence 34, Application US/10311453  
; Publication No. US20040031073A1  
; GENERAL INFORMATION:  
; APPLICANT: Werner, Thomas  
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and  
; FILE REFERENCE: 1226-4  
; CURRENT APPLICATION NUMBER: US/10/311,453  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: EP 00870132.8  
; PRIOR FILING DATE: 2000-06-16  
; PRIOR APPLICATION NUMBER: US 60/258,415  
; PRIOR FILING DATE: 2000-12-27  
; PRIOR APPLICATION NUMBER: EP 01870053.4  
; PRIOR FILING DATE: 2001-03-16  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 34  
; LENGTH: 1620  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-10-014-101-34  
Alignment Scores:  
Pred. No.: 539 Length: 1620  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 14 Indels: 0  
DB: 14 Gaps: 0  
US-10-014-101B-39 (1-7) x US-10-014-101-34 (1-1620)  
QY 1 ValGlyGlyThrLeuSerAan 7  
DB 508 GTTGGAGGTACACTCTCCAAT 528  
US-10-311-453-34  
; Sequence 34, Application US/10311453  
; Publication No. US20040031073A1  
; GENERAL INFORMATION:  
; APPLICANT: Werner, Thomas  
; TITLE OF INVENTION: Method for modifying plant morphology, biochemistry and  
; FILE REFERENCE: 1226-4  
; CURRENT APPLICATION NUMBER: US/10/311,453  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: EP 00870132.8  
; PRIOR FILING DATE: 2000-06-16  
; PRIOR APPLICATION NUMBER: US 60/258,415  
; PRIOR FILING DATE: 2000-12-27  
; PRIOR APPLICATION NUMBER: EP 01870053.4  
; PRIOR FILING DATE: 2001-03-16  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 34  
; LENGTH: 1620  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-10-311-453-34  
Alignment Scores:  
Pred. No.: 539 Length: 1620  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 14 Indels: 0  
DB: 17 Gaps: 0  
US-10-014-101B-39 (1-7) x US-10-311-453-34 (1-1620)  
QY 1 ValGlyGlyThrLeuSerAan 7  
DB 508 GTTGGAGGTACACTCTCCAAT 528

RESULT 36  
US-10-326-184-12  
; Sequence 12, Application US/10326184  
; Publication No. US20030163847A1  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Company  
; TITLE OF INVENTION: REVERSIBLE MALE STERILITY IN TRANSGENIC PLANTS BY EXPRESSION OF C  
; FILE REFERENCE: MTC6781.1  
; CURRENT APPLICATION NUMBER: US/10/326,184  
; CURRENT FILING DATE: 2002-12-20  
; PRIOR APPLICATION NUMBER: US 60/343,129  
; PRIOR FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 12  
; LENGTH: 1623  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-10-326-184-12  
Alignment Scores:  
Pred. No.: 540 Length: 1623  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 16 Gaps: 0  
US-10-014-101B-39 (1-7) x US-10-326-184-12 (1-1623)  
QY 1 ValGlyGlyThrLeuSerAan 7  
DB 511 GTTGGAGGTACACTCTCCAAT 531  
US-10-817-483-32  
; Sequence 32, Application US/10817483  
; Publication No. US20040237147A1  
; GENERAL INFORMATION:  
; APPLICANT: Habben, Jeffrey E.  
; APPLICANT: Zinsemeier, Christopher  
; APPLICANT: Tomes, Dwight  
; APPLICANT: Abbott, Shane  
; APPLICANT: Helentjaris, Timothy G.  
; APPLICANT: Niu, Xiaomu  
; TITLE OF INVENTION: Modulation of Cytokinin Activity in  
; FILE REFERENCE: 0803R  
; CURRENT APPLICATION NUMBER: US/10/817,483  
; CURRENT FILING DATE: 2004-04-02  
; PRIOR APPLICATION NUMBER: US 60/460,718  
; PRIOR FILING DATE: 2003-04-04  
; PRIOR APPLICATION NUMBER: US 09/545,334  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: US 60/129,844  
; PRIOR FILING DATE: 1999-04-16  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 32  
; LENGTH: 1629  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(1629)  
US-10-817-483-32  
Alignment Scores:  
Pred. No.: 542 Length: 1629  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservatives: 0

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 18 Gaps: 0

US-10-014-101B-39 (1-7) x US-10-817-483-32 (1-1629)

QY 1 ValGlyGlyThrLeuSerAsn 7

DB 502 GTCGGCGGACGCTGTCCAAT 522

RESULT 38

US-10-326-184-20

; Sequence 20, Application US/10326184

; Publication No. US20030163847A1

; GENERAL INFORMATION:

; APPLICANT: Monsanto Company

; TITLE OF INVENTION: REVERSIBLE MALE STERILITY IN TRANSGENIC PLANTS BY EXPRESSION OF

; TITLE OF INVENTION: OXIDASE 1

; FILE REFERENCE: WTC6781.1

; CURRENT APPLICATION NUMBER: US/10/326,184

; CURRENT FILING DATE: 2002-12-20

; PRIOR APPLICATION NUMBER: US 60/343,129

; PRIOR FILING DATE: 2001-12-20

; NUMBER OF SEQ ID NOS: 24

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 20

; LENGTH: 1655

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-10-326-184-20

Alignment Scores:

Pred. No.: 551 Length: 1655  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 16 Gaps: 0

US-10-014-101B-39 (1-7) x US-10-326-184-20 (1-1655)

QY 1 ValGlyGlyThrLeuSerAsn 7

DB 535 GTCGGCGGACGCTGTCCAAT 555

RESULT 39

US-10-437-963-70571

; Sequence 70571, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 70571

; LENGTH: 1670

; TYPE: DNA

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_7112C.1

US-10-437-963-70571

Alignment Scores:

Pred. No.: 556 Length: 1670  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 18 Gaps: 0

US-10-014-101B-39 (1-7) x US-10-437-963-70571 (1-1670)

QY 1 ValGlyGlyThrLeuSerAsn 7

DB 603 GTCGGCGGACGCTGTCCAAT 623

RESULT 40

US-10-437-963-83827

; Sequence 83827, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 83827

; LENGTH: 1673

; TYPE: DNA

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_83121C.1

US-10-437-963-83827

Alignment Scores:

Pred. No.: 557 Length: 1673  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 18 Gaps: 0

US-10-014-101B-39 (1-7) x US-10-437-963-83827 (1-1673)

QY 1 ValGlyGlyThrLeuSerAsn 7

DB 609 GTCGGCGGACGCTGTCCAAT 629

Search completed: February 18, 2005, 10:12:20  
Job time : 333.98 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 18, 2005, 02:20:24 ; Search time 825.44 Seconds  
(without alignments)  
322.797 Million cell updates/sec

Title: US-10-014-101B-39

Perfect score: 35

Sequence: 1 VGGTSLN 7

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10014101/runat\_16022005\_075846\_7952/app.query.fasta\_1.796  
-DB=EST -OPMT=fastap -SUFFIX=ret -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bites -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL  
-OUTFMT=ptco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10014101 @CGN 1.1 5453 @runat\_16022005\_075846\_7952 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DSV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:  
1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_hic.\*  
4: gb\_est3.\*  
5: gb\_est4.\*  
6: gb\_est5.\*  
7: gb\_est6.\*  
8: gb\_gss1.\*  
9: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	100.0	116	9	CG802105 1118027F1
2	35	100.0	162	9	BX890851 Arabidops
3	35	100.0	230	8	BZ643095 OGAP23TC
4	35	100.0	299	9	CR490530 Medicago
5	35	100.0	306	1	AV424619 AV424619
6	35	100.0	307	8	CC131284 NDL.32E23
7	35	100.0	309	5	BX476387 DKFZp686N
8	35	100.0	362	8	BZ622937 ig81e07.b
9	35	100.0	399	9	CE648629 tigr-gss-

C	10	35	100.0	445	6	CD703115	EST19706
	11	35	100.0	451	5	BQ588342	E012308-0
	12	35	100.0	461	1	AU289504	AU289504
C	13	35	100.0	470	8	AQ795075	AQ795075 nbxb0055C
	14	35	100.0	473	1	AU243228	AU243228
	15	35	100.0	476	4	B1930687	EST550576
	16	35	100.0	481	4	BMI37409	WHE0463-0
	17	35	100.0	485	9	CG927858	MBE1159TR
	18	35	100.0	487	5	BP031101	BP031101
C	19	35	100.0	491	9	CR399680	CR399680 Arabidops
	20	35	100.0	504	9	CC656142	OGUD893TH
	21	35	100.0	506	8	AQ827482	HS_5304_A
	22	35	100.0	511	6	CB640717	OSJNea16A
C	23	35	100.0	515	5	BQ869165	BQ869165 QSD5F19.Y
	24	35	100.0	518	5	BX283649	BX283649
C	25	35	100.0	520	7	CR775437	DKFZp469C
	26	35	100.0	525	4	BI498566	sa115g12
C	27	35	100.0	536	8	AQ510474	nbxb0095B
	28	35	100.0	540	2	AW224503	EST302945
	29	35	100.0	542	7	CF603749	BACCA01_0
	30	35	100.0	550	9	CNS000MU	AL083412 Arabidops
	31	35	100.0	551	9	CG357974	CG357974 OZBGS58TV
C	32	35	100.0	556	7	CV057457	BNEL28a6
	33	35	100.0	559	4	BI700658	sa127e03
	34	35	100.0	564	2	AW736360	AW736360 EST332279
C	35	35	100.0	566	8	CC022210	CC022210 3591_1_28
	36	35	100.0	567	5	BU084470	BU084470
C	37	35	100.0	567	9	CL157422	CL157422 104_345_1
C	38	35	100.0	571	8	BZ706895	BZ706895 OGBD30TC
C	39	35	100.0	573	7	CO600263	DGB-199k1
	40	35	100.0	574	5	BP133582	BP133582
	41	35	100.0	576	8	BZ949154	BZ949154 CH240_37E
C	42	35	100.0	583	7	CV057888	CV057888 BNEL31h8
	43	35	100.0	584	4	BGI30984	BGI30984 EST463876
	44	35	100.0	584	8	BZ333618	BZ333618 hx71d05_9
	45	35	100.0	585	2	BF648633	BF648633 NF048D04E

#### ALIGNMENTS

CG802105 1118027F12.y1 1118 - RescueMu Grid S Zea maye genomic, genomic  
survey sequence.

CG802105

CG802105.1 GI:38237891

GSS.

Zeas maye

Zeas maye

Zeas maye

Zeas maye

Zeas maye

Zeas maye

Zeas maye

Zeas maye

Zeas maye

Zeas maye

Zeas maye

Zeas maye

Zeas maye

Zeas maye

Zeas maye

Zeas maye

Zeas maye

Zeas maye

Zeas maye

Zeas maye

Zeas maye

Zeas maye

Zeas maye

Zeas maye

Zeas maye

Zeas maye

Zeas maye

Zeas maye

Zeas maye

```

/db_xref="taxon:4577"
/tissue_type="leaf"
/dev stage="adult"
/lab_host="DH10B"
/clone_lib="1118 - RescueMu Grid S"
/note="Organ: leaf; Vector: RescueMu (engineered from pBlueScript backbone); Site 1: BamHI; Site 2: BglII; RescueMu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site 'www.zmbd.iastate.edu' and follow the links for 'RescueMu.' Grid S was grown at San Diego in 2002. DNA was extracted from leaf strips, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

ORIGIN
Alignment Scores:
Pred. No.: 225 Length: 116
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-014-101B-39 (1-7) x CG802105 (1-116)

QY 1 ValGlyGlyThrLeuSerAsn 7
Db 61 GTCGGGGCACACTGTCCAAC 91

RESULT 2
LOCUS BX890851 162 bp DNA linear GSS 05-APR-2004
DEFINITION Arabidopsis thaliana T-DNA flanking sequence GK-157A04-013241, genomic survey sequence.
ACCESSION BX890851
VERSION BX890851.1 GI:39923346
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
Li, Y., Rosso, M.G., Strizhov, N., Viehoveer, P., and Weissshaar, B. GABI-Kat SimpleSearch, a flanking sequence tag (FST) database for the identification of T-DNA insertion mutants in Arabidopsis thaliana
Bioinformatics 19 (11), 1441-1442 (2003)
MEDLINE 22755829
PUBMED 12874060
REFERENCE 2
AUTHORS Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K. and Weissshaar, B.
TITLE An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics
JOURNAL Plant Mol. Biol. 53 (1-2), 247-259 (2003)
MEDLINE 23117147
PUBMED 14756321
REFERENCE 3
AUTHORS Strizhov, N., Li, Y., Rosso, M.G., Viehoveer, P., Dekker, K.A. and Weissshaar, B.
TITLE High-throughput generation of sequence indexes from T-DNA mutagenized Arabidopsis thaliana lines
JOURNAL Biotechniques 35 (6), 1164-1168 (2003)
PUBMED 14682050
REFERENCE 4 (bases 1 to 162)
AUTHORS Li, Y., Strizhov, N., Rosso, M.G. and Weissshaar, B.
TITLE Direct Submission
JOURNAL Submitted (31-MAR-2004) Weissshaar B., Max-Planck-Institut fuer

```

Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany  
It indicates an insertion close to or within gene At5G56970.  
Details on the protocols used for generation of the sequence are described in References 1-3. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-Kat project.  
'GABI-Kat' is part of the German Plant Genomics program designated 'GABI'. Information on line availability can be found at:  
http://www.mpiz-koeln.mpg.de/GABI-Kat/

Location/Qualifiers  
1..162  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/strain="Columbia 0"  
/db\_xref="taxon:3702"  
/clone="GK-157A04-013241"  
/clone\_lib="Arabidopsis thaliana T-DNA insertion lines"  
/ecotype="Col-0"  
/note="PCR was performed on DNA from Arabidopsis thaliana plants (T1) which were transformed with the T-DNA from vector PAC161 (GenBank accession number: AJ537514). The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed."

ORIGIN
Alignment Scores:
Pred. No.: 321 Length: 162
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-014-101B-39 (1-7) x BX890851 (1-162)

QY 1 ValGlyGlyThrLeuSerAsn 7
Db 71 GTCGGTGGGACGTATCAAC 91

RESULT 3
BZ643095/230 bp DNA linear GSS 29-JAN-2003
LOCUS OGAOF23TC ZM 0.7 1.5 KB Zea mays genomic clone ZMWBMa0107C22, genomic survey sequence.
DEFINITION BZ643095
ACCESSION BZ643095
VERSION BZ643095.1 GI:28104206
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 230)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N. Consortium for Maize Genomics
Unpublished (2002)
Other\_GSSs: OGAOF23TM
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@cigr.org
Seq primer: Tg
Class: sheared ends.
Location/Qualifiers
1..230
/organism="Zea mays"
/mol\_type="genomic DNA"

/strain="B73"  
 /db\_xref="taxon:4577"  
 /clone="ZMBMa0107C22"  
 /clone\_lib="ZM 0.7 1.5 KB"  
 /notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb  
 methylation filtered genomic DNA library"

## ORIGIN

Alignment Scores:  
 Pred. No.: 465 Length: 230  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 8 Gaps: 0

US-10-014-101B-39 (1-7) x BZ643095 (1-230)

QY 1 ValGlyGlyThrLeuSerAsn 7  
 |||||  
 DB 80 GTCGGGGCACGCTGTCCAC 60

## RESULT 4

CR490530/c  
 LOCUS  
 DEFINITION Medicago truncatula BAC ends cultivar Jemalong A17 of Medicago truncatula, genomic survey sequence.

ACCESSION CR490530.1 GI:48652106

KEYWORDS GSS.

SOURCE Medicago truncatula (barrel medic)

ORGANISM Medicago truncatula  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
 Medicago.

## REFERENCE

1 (bases 1 to 299)  
 Genoscope.  
 TITLE Direct Submission  
 JOURNAL Submitted (10-JUN-2004) Genoscope - Centre National de Sequencage :  
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
 - Web : www.genoscope.cns.fr)

## FEATURES

source  
 1..299  
 Location/Qualifiers  
 /organism="Medicago truncatula"  
 /mol\_type="genomic DNA"  
 /cultiivar="Jemalong A17"  
 /db\_xref="taxon:3880"  
 /clone\_lib="MTH2"  
 /notes="Vector: pBelOBAC11 ; Site 1: HindIII ; Site 2:  
 HindIII ; Cook, D.R. and Kim, D.J-Genoscope sequence ID :  
 mth2-161F3FM1"

## ORIGIN

Alignment Scores:  
 Pred. No.: 613 Length: 299  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-10-014-101B-39 (1-7) x CR490530 (1-299)

QY 1 ValGlyGlyThrLeuSerAsn 7  
 |||||  
 DB 42 GTTGGAGGCACTCTTCGAAT 22

## RESULT 5

AV424619  
 LOCUS  
 DEFINITION AV424619 Lotus japonicus young plants (two-week old) Lotus  
 corniculatus var. japonicus cDNA clone MWM042d03\_r 5', mRNA

## sequence.

AV424619  
 VERSION AV424619.1 GI:7781717  
 EST.  
 SOURCE  
 ORGANISM

Lotus corniculatus var. japonicus (Lotus japonicus)

Lotus corniculatus var. japonicus  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;  
 Lotus.

REFERENCE 1 (bases 1 to 306)

Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.

AUTHORS Generation of 7137 non-redundant expressed sequence tags from a

TITLE legume, Lotus japonicus

JOURNAL DNA Res. 7 (2), 127-130 (2000)

MEDLINE 20277479

PUBMED 10819328

COMMENT Contact: Yasukazu Nakamura

The First Laboratory for Plant Gene Research

Kazusa DNA Research Institute

Yana 1532-3, Kisarazu, Chiba 292-0812, Japan

Email: ynakamu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

## FEATURES

Location/Qualifiers  
 1..306  
 /organism="Lotus corniculatus var. japonicus"  
 /mol\_type="mRNA"  
 /isolate="Miyakojima MG-20"  
 /db\_xref="taxon:34305"  
 /clone="MWM042d03\_r"  
 /tissue\_type="whole plant"  
 /dev\_stage="two-week old plant"  
 /clone\_lib="Lotus japonicus young plants (two-week old)"  
 /note="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:  
 XhoI; synonym: Lotus japonicus"

## ORIGIN

Alignment Scores:  
 Pred. No.: 629 Length: 306  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 1 Gaps: 0

US-10-014-101B-39 (1-7) x AV424619 (1-306)

QY 1 ValGlyGlyThrLeuSerAsn 7  
 |||||  
 DB 173 GTGGGAGGACGTTATCTAAC 193

## RESULT 6

CC131284/c  
 LOCUS  
 DEFINITION ND.L.32E23.T7 Notre Dame Liverpool Aedes aegypti genomic clone  
 ND.L.32E23, genomic survey sequence.

ACCESSION CC131284

VERSION CC131284.1 GI:30000339

KEYWORDS GSS.

SOURCE Aedes aegypti (yellow fever mosquito)

ORGANISM

Aedes aegypti  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Aedes;  
 Stegomyia.

REFERENCE 1 (bases 1 to 307)

AUTHORS Loftus, B., Shetty, J., Knudson, D. and Severson, D.

TITLE BAC end sequencing of Aedes aegypti

JOURNAL Unpublished (2003)

COMMENT Other GSSs: ND.L.32E23.SP6

Contact: Brendan Loftus

Department of Eukaryotic Genomics

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-3543

Fax: 301-838-0208  
Email: enta@igr.org  
Library was provided by David Severson  
Seq primer: T7  
Class: BAC ends

# FEATURES

source  
Location/Qualifiers  
1..307  
/organism="Aedes aegypti"  
/mol\_type="genomic DNA"  
/strain="liverpool"  
/db\_xref="taxon:7159"  
/clone="NDL32E23"  
/clone\_lib="Notre Dame Liverpool"  
/notes="Vector: pECBAC1; Site 1: Hind III; The library was prepared from whole body tissue of newly hatched L1 larvae by David Severson at the University of Notre Dame and Hongbin Zhang"

# ORIGIN

Alignment Scores:  
Pred. No.: 631 Length: 307  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0

US-10-014-101B-39 (1-7) x CC131284 (1-307)

QY 1 ValGlyGlyThrLeuSerAsn 7  
Db 98 GTCGGTGGTACACTGTCCAAC 78

# RESULT 7

US-10-014-101B-39 (1-7) x CC131284 (1-307)

LOCUS BX476387/c 309 bp mRNA linear EST 04-SEP-2003  
DEFINITION DKFP686N06185\_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone  
ACCESSION DKFP686N06185  
VERSION BX476387  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 309)  
AUTHORS Bloembergen, H., Boecher, M., Mewes, H.W., Weil, B., Amid, C., Oeanger, A., Fobo, G., Han, M., and Wiemann, S.  
TITLE EST (Bloembergen, H., Boecher, M., Mewes, H.W., Weil, B., Amid, C., et al.)  
JOURNAL Unpublished (2003)  
COMMENT Contact: MIPS

MIPS  
Ingolstaedter Landstr. 1, D-85764 Neuherberg, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;  
Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.  
No 5' sequence available.

This clone (DKFP686N06185) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de.

# FEATURES

source  
Location/Qualifiers  
1..309  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFP686N06185"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="686 (synonym: hlcc3)"  
/note="Vector: pTriplex2; Site\_1: SfiI; Site\_2: SfiIB;

# ORIGIN

Alignment Scores:  
Pred. No.: 635 Length: 309  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 5 Gaps: 0

US-10-014-101B-39 (1-7) x BX476387 (1-309)

QY 1 ValGlyGlyThrLeuSerAsn 7  
Db 108 GTGGCGGACGCTTCTTAAT 88

# RESULT 8

US-10-014-101B-39 (1-7) x BX476387 (1-309)

LOCUS BZ622937/c 362 bp DNA linear GSS 16-JAN-2003  
DEFINITION ig81e07\_b1 WGS-ZmaysF (DH5a methyl filtered) Zea mays genomic clone  
ACCESSION BZ622937  
VERSION BZ622937.1 GI:27774789  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays

REFERENCE 1 (bases 1 to 362)  
AUTHORS Rabinowicz, P.D., O'Shaughnessy, A.L., Balija, V., Dedhia, N., Katzenburger, P., King, L., Miller, B., Muller, S., Nascimento, L., Zutavern, T., McCombie, W.R., and Martienssen, R.A.

Genomic shotgun sequences from Zea mays (methyl-filtered)  
Unpublished (2002)  
Contact: W. Richard McCombie  
Lita Annenberg Hazen Genome Sequencing Center  
Cold Spring Harbor Laboratory  
PO Box 100, Cold Spring Harbor, NY 11724, USA  
Tel: 516 367 8884  
Fax: 516 367 8874  
Email: mcombie@cshl.org  
Plate: ig81 row: e column: 07  
Seq primer: -21M13UnivFwd  
Class: shotgun  
High quality sequence stop: 362.  
Location/Qualifiers  
1..362  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/cultivar="B73"  
/db\_xref="taxon:4577"  
/clone="ig81e07"  
/lab\_host="DH5a"  
/clone\_lib="WGS-ZmaysF (DH5a methyl filtered)"  
/note="Organ: immature ears; Site 1: Xba I; Site 2: Xba I; The vector was digested with XbaI and one nucleotide was added by fill in in the recessive 3' end. The genomic DNA was nebulized, end repaired, adaptor ligated and size fractionated using sephadex. The resulting fragments were between 0.8 and 3 kb and were cloned into the vector (.x/y reads in M13mp19, .b/g reads in pUC19). The same ligation was transformed into DH5a."

# JOURNAL

COMMENT

# FEATURES

source

Alignment Scores:  
Pred. No.: 751 Length: 362  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0

US-10-014-101B-39 (1-7) x BZ622937 (1-362)

Qy 1 ValGlyGlyThrLeuSerAsn 7  
Db 136 GTGGGGGACGCTCTCCAC 116

RESULT 9

CE648629

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

Length:

Matches:

Conservative:

Mismatches:

Indels:

Gaps:

US-10-014-101B-39 (1-7) x CE648629 (1-399)

Qy 1 ValGlyGlyThrLeuSerAsn 7

Db 127 GTTGGAGGACCTTTCAC 147

RESULT 10

CD703115/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

Length:

Matches:

Conservative:

Mismatches:

Indels:

Gaps:

US-10-014-101B-39 (1-7) x CE648629 (1-399)

Qy 1 ValGlyGlyThrLeuSerAsn 7

Db 127 GTTGGAGGACCTTTCAC 147

RESULT 10

CD703115/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

US-10-014-101B-39 (1-7) x BZ622937 (1-362)

Qy 1 ValGlyGlyThrLeuSerAsn 7

Db 136 GTGGGGGACGCTCTCCAC 116

RESULT 9

CE648629

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

Length:

Matches:

Conservative:

Mismatches:

Indels:

Gaps:

US-10-014-101B-39 (1-7) x CE648629 (1-399)

Qy 1 ValGlyGlyThrLeuSerAsn 7

Db 127 GTTGGAGGACCTTTCAC 147

RESULT 10

CD703115/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

Length:

Matches:

Conservative:

Mismatches:

Indels:

Gaps:

US-10-014-101B-39 (1-7) x CE648629 (1-399)

Qy 1 ValGlyGlyThrLeuSerAsn 7

Db 127 GTTGGAGGACCTTTCAC 147

RESULT 10

CD703115/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Zeng, Y.-X.

Transcriptional Gene Expression Profile of Human Nasopharynx

Unpublished (2003)

Contact: Yixin Zeng

Cancer Center

Sun Yat-sen University

651 Dongfeng Road East, Guangzhou 510060, China

Tel: 86-1380-9770-743

Fax: 86-20-8775-4506

Email: yxzeng@zsus.edu.cn.

Location/Qualifiers

1..445

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/tissue type="normal nasopharynx"

/clone\_lib="human nasopharynx"

/note="ESTs generated from a normal nasopharynx cDNA library from southern Chinese"

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

Length:

Matches:

Conservative:

Mismatches:

Indels:

Gaps:

US-10-014-101B-39 (1-7) x CD703115 (1-445)

Qy 1 ValGlyGlyThrLeuSerAsn 7

Db 111 GTGGGGGACGCTTTCAT 91

RESULT 11

BQ588342

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Weisshaar B

ADIS DNA core facility at MPIZ

Max-Planck-Institute for Plant Breeding Research

Carl-von-Linne Weg 10, 50829 Koeln, Germany

Fax: 00492215062851

Email: weisshaar@piz-koeln.mpg.de

Insert Length: 451 Std Error: 0.00

Plate: 8 row: P column: 04

Seq primer: SP6; CATACGATTAGTGACACTATAG.

Location/Qualifiers

1..451

/organism="Beta vulgaris"

/mol\_type="mRNA"

/cultivar="KWS2320 (double haploid, monogerm breeding line)"

/db\_xref="GABI:184499"

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

Length:

Matches:

Conservative:

Mismatches:

Indels:

Gaps:

US-10-014-101B-39 (1-7) x BQ588342 (1-451)

Qy 1 ValGlyGlyThrLeuSerAsn 7

Db 111 GTGGGGGACGCTTTCAT 91

RESULT 11

BQ588342

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Weisshaar B

ADIS DNA core facility at MPIZ

Max-Planck-Institute for Plant Breeding Research

Carl-von-Linne Weg 10, 50829 Koeln, Germany

Fax: 00492215062851

Email: weisshaar@piz-koeln.mpg.de

Insert Length: 451 Std Error: 0.00

Plate: 8 row: P column: 04

Seq primer: SP6; CATACGATTAGTGACACTATAG.

Location/Qualifiers

1..451

/organism="Beta vulgaris"

/mol\_type="mRNA"

/cultivar="KWS2320 (double haploid, monogerm breeding line)"

/db\_xref="GABI:184499"

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

Length:

Matches:

Conservative:

Mismatches:

Indels:

Gaps:

US-10-014-101B-39 (1-7) x BQ588342 (1-451)

Qy 1 ValGlyGlyThrLeuSerAsn 7

Db 111 GTGGGGGACGCTTTCAT 91

RESULT 11

BQ588342

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

```

/db_xref="taxon:161934"
/clone="024-008-P04"
/tissue_type="leaf"
/lab_host="EMDH108"
/clone_lib="WP1Z-ADIS-024-leaf"
/notes="Vector: pCMVSPORT6; Site 1: SalI; Site 2: NotI;
cDNA library from sugar beet, library provided by KWS
Kleinwanzlebener Saatgut AG Einbeck, Germany, contact:
b.schulz@kws.de; cloning sites SalI-NotI, primer sites and
orientation:
SP6-Sali-CCACGCGTCG-5prime-cDNA-polyA-CC-NotI-T7; Note:
Sequencing granted in the context of the GABI-Beet
project, local PI: Dr. Katharina Schneider, coordinator:
Prof. Christian Jung; Sequence submission managed by
RZPD/GABI-Primary database:http://gabi.rzpd.de"

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 948 Length: 451
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservatative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

```

US-10-014-101B-39 (1-7) x BQ588342 (1-451)

```

Qy 1 ValGlyGlyThrLeuSerAsn 7
Db 152 GTGGTGGTACTTGTCTAT 172

```

## RESULT 12

## AU289504

## LOCUS

DEFINITION AU289504 zinnia cultured mesophyll cell equalized cDNA Zinnia

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Asterales; Asteraceae; Asteroideae; Heliantheae; Zinnia.

## REFERENCE

## AUTHORS

Demura, T., Tashiro, G., Horiguchi, G., Kishimoto, N., Kubo, M., Matsuo, N., Minami, A., Nagata-Hiwatashi, M., Nakamura, K., Okamura, Y., Sassa, N., Suzuki, S., Yazaki, J., Kikuchi, S. and Fukuda, H.

Visualization by comprehensive microarray analysis of gene expression programs during transdifferentiation of mesophyll cells into xylem cells

## JOURNAL

## COMMENT

Contact: Taku Demura

Morphogenesis Research Group

RKEN Plant Science Center

1-7-22 Suehirocho, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9605

Fax: 81-45-503-9573

Email: demura@postman.riken.go.jp

This clone was obtained at our laboratory.

Seq primer: M13 forward.

## FEATURES

## source

1. 461

/organism="Zinnia elegans"

/mol\_type="mRNA"

/cultivar="Canary bird"

/db\_xref="taxon:34245"

/clone="24008"

/tissue\_type="mesophyll cell"

/clone\_lib="zinnia cultured mesophyll cell equalized cDNA"

/notes="Vector: pGEM-T easy; cultured in tracheary element

## ORIGIN

differentiation-inductive medium"

## Alignment Scores:

```

Pred. No.: 970 Length: 461
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservatative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

```

US-10-014-101B-39 (1-7) x AU289504 (1-461)

```

Qy 1 ValGlyGlyThrLeuSerAsn 7
Db 384 GTAGGTGGTACTTGTCTAAC 404

```

## RESULT 13

## AQ795075/c

## LOCUS

DEFINITION AQ795075 470 bp DNA linear GSS 04-AUG-1999  
nbxb0055C05r CUGI Rice BAC Library Oryza sativa (japonica cultivar-group) genomic clone nbxb0055C05r, genomic survey sequence.

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

Oryza sativa (japonica cultivar-group)  
Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

Wing, R.A. and Dean, R.A.  
A BAC End Sequencing Framework to Sequence the Rice Genome

Unpublished (1998)

Contact: Wing RA

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288

Fax: 864 656 4293

Email: rwing@clemson.edu

Seq primer: GGAAACAGCTATGACCATG

Class: BAC ends

High quality sequence stop: 319.

Location/Qualifiers

1. 470

/organism="Oryza sativa (japonica cultivar-group)"

/mol\_type="genomic DNA"

/cultivar="Nipponbare"

/db\_xref="taxon:39947"

/clone="nbxb0055C05r"

/tissue\_type="Leaf"

/lab\_host="E. coli DH10B"

/clone\_lib="CUGI Rice BAC Library"

/note="Vector: pBeloBAC11; Site 1: HindIII; Site 2:

HindIII; Rice is one of two most popular grains in the

world. Half of the world population especially those

inhabiting highly populated areas of the humid tropics

and subtropics, rely on rice as their primary source of

carbohydrate. Monocotyledonous rice is a diploid plant

(2n=24) with a haploid genome equivalent of 431 Mbp

(Arumuganathan and Earle, 1991). The relatively small

genome of rice, three times larger than that of

Arabidopsis, makes it suitable for genomic studies. In

order to facilitate positional cloning, physical mapping

and genome sequencing of rice, we have constructed a BAC

library from Oryza sativa, Nipponbare variety. The

library contains 36,864 clones with an average insert size

of 128.5 Kb providing 10.9 haploid genome equivalents. The

deep coverage allows the isolation a particular sequence

with a probability of 99.9 %. Two high density filters,

each containing 18,432 clones (doubly spotted), represent

the whole library for colony screening."

```

ORIGIN
Alignment Scores:
Pred. No.: 991 Length: 470
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservatative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0

US-10-014-101B-39 (1-7) x AQ795075 (1-470)

QY 1 ValGlyGlyThrLeuSerAsn 7
Db 417 GTTGGAGGCACACTGTGCAAT 397

RESULT 14
LOCUS AU243228 473 bp mRNA linear EST 15-JAN-2002
DEFINITION AU243228 Medaka eye cDNA library (SNK01) Oryzias latipes cDNA clone
ACCESSION NGY29.03f, mRNA sequence.
VERSION AU243228
KEYWORDS AU243228.1 GI:18155807
SOURCE EST.
ORGANISM Oryzias latipes (Japanese medaka)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
1 (bases 1 to 473)
Sanaka, E., Hori, H., Naruse, K., Mitani, H. and Tanaka, M.
Medaka EST analysis
Unpublished (2001)
Contact: Emi Sanaka
Department of Biological Sciences
Graduate School of Science, Nagoya University
Furo-cho, Chikusa-ku, Nagoya 464-8602, Japan
Tel: 81-52-789-2973
Fax: 81-52-789-2974
Email: sanaka@bio.nagoya-u.ac.jp
This clone was isolated from Medaka eye cDNA library (SNK01) 5'end.

FEATURES
source
Location/Qualifiers
1..473
/organism="Oryzias latipes"
/mol type="mRNA"
/strain="wild type"
/db_xref="taxon:8090"
/clone="NGY29.03f"
/tissue type="eye"
/dev stage="adult"
/clone lib="Medaka eye cDNA library (SNK01)"
/note="Wild samples from Okayama Pref. (Southern part of Japan)"

ORIGIN
Alignment Scores:
Pred. No.: 997 Length: 473
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservatative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-10-014-101B-39 (1-7) x AU243228 (1-473)

QY 1 ValGlyGlyThrLeuSerAsn 7
Db 446 GTTGGGGGACACTTAGCAAT 466

RESULT 15
B1930687

```

```

LOCUS B1930687 476 bp mRNA linear EST 18-OCT-2001
DEFINITION EST550576 tomato flower, 8 mm to preanthesis buds Lycopersicon
esculentum cDNA clone cTOC17F5 5' end, mRNA sequence.
ACCESSION B1930687
VERSION B1930687.1 GI:16245159
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 476)
van der Hoeven, R.S., Bezzerides, J.L., Karamycheva, S.A., Tsai, J.,
Utterback, T., Van Aken, S., Ronning, C.M., Nierman, W., Fraser, C.M.,
Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.
Generation of ESTs from tomato flower tissue, buds 8 mm -
preanthesis
Unpublished (2001)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
This clone is available through the Clemson University Genomics
Institute
Seq primer: T3.
Location/Qualifiers
1..476
/organism="Lycopersicon esculentum"
/mol type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cTOC17F5"
/tissue type="flower"
/dev stage="buds 8mm to preanthesis"
/clone lib="tomato flower, 8 mm to preanthesis buds"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Cornell University; sequencing: The
Institute for Genomic Research; Flower buds and flowers
were taken from greenhouse plants (4-8 wks old, TA496).
They were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."

ORIGIN
Alignment Scores:
Pred. No.: 1e+03 Length: 476
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservatative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-014-101B-39 (1-7) x B1930687 (1-476)

QY 1 ValGlyGlyThrLeuSerAsn 7
Db 281 GTCGGAGGTACTCTGTCTAAT 301

RESULT 16
B1937409
LOCUS B1937409 481 bp mRNA linear EST 28-NOV-2001
DEFINITION WHE0463-0466_K01_K012s Wheat Fusarium graminearum infected spike
cDNA library Triticum aestivum cDNA clone WHE0463-0466_K01_K01,
mRNA sequence.
ACCESSION B1937409
VERSION B1937409.1 GI:17146176
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
1 (bases 1 to 481)

```

**AUTHORS** Anderson,O.D., Chao,S., Han,P.S., Han,P.S., Heinen,S., Hsia,C.C., Kang,Y., Kruger,W.M., Lazo,G.R., Miller,S., Muehlbauer,G.J., Miller,R., Pritsch,C., Rausch,C.J., Seaton,C.L., Tong,J.C., Vance,C. and Wilson,C.F.

**TITLE** The structure and function of the expressed portion of the wheat genomes - Fusarium graminearum infected spike cDNA library

**JOURNAL** Unpublished (2001)

**COMMENT** Contact: Olin Anderson  
US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center  
800 Buchanan Street, Albany, CA 94710, USA  
Tel: 5105595773  
Fax: 5105595818  
Email: oanderson@pw.usda.gov

Sequences have been trimmed to remove vector sequence and low quality sequence with phred score less than 20. No effort was taken to identify ESTs of fungal origin from this library, thus this EST could be of wheat or fungal origin.

Seq primer: Stratagene SK primer.

# FEATURES

source  
1. .481  
Location/Qualifiers  
/organism="Triticum aestivum"  
/mol\_type="mRNA"  
/cultivar="Sumai3"  
/db\_xref="taxon:4565"  
/clone="WHE0463-0466 K01\_K01"  
/tissue\_type="Spike"  
/dev\_stage="Adult plant"  
/lab\_host="E. coli SOL5"  
/clone\_lib="Wheat Fusarium graminearum infected spike cDNA library"  
/notes="Vector: Lambda Uni-ZAP XR, excised phagemid pBluescript SK; Site 1: EcoRI; Site 2: XhoI; Plants were grown in the greenhouse. Spikes were sprayed at anthesis with Fusarium graminearum. Total RNA, and poly(A) RNA were prepared and pooled from infected spike at 0, 6, 12, 24, 36 and 48 hours after inoculation, a cDNA library was made, and the cDNA clones were in vivo excised to give pBluescript phagemids in G. Muehlbauer lab at the University of Minnesota (Kruger, W.M., Muehlbauer, G.J., Pritsch, C., Vance, C.). The cDNA library should contain genes of both wheat and fungal pathogen origin. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

# ORIGIN

Alignment Scores:  
Pred. No.: 1.02e+03 Length: 481  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-39 (1-7) x BM137409 (1-481)

QY 1 ValGlyGlyThrLeuSerAsn 7  
|||||  
Db 205 GTGGGGGCGACCTCTCCAC 225

# RESULT 17

CG927858  
LOCUS MBEII597R mth2 Medicago truncatula genomic clone 62121, genomic survey sequence.  
DEFINITION CG927858  
ACCESSION CG927858.1 GI:39788945  
VERSION  
KEYWORDS  
SOURCE Medicago truncatula (barrel medic)

ORGANISM  
Medicago truncatula  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;

# REFERENCE

1 (bases 1 to 485)  
TOWN,C.D., Shetty,J., Koo,H. and Feldblyum,T.F.  
Sequencing of BAC ends from Medicago truncatula Unpublished (2003)  
Other GSSs: MBEII597F  
Contact: Chris Town

# TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
Seq primer: CAGGAACAGCTATGACC  
Class: BAC ends.

# FEATURES

source  
1. .485  
Location/Qualifiers  
/organism="Medicago truncatula"  
/mol\_type="genomic DNA"  
/cultivar="genotype Al7"  
/db\_xref="taxon:3880"  
/clone="62121"  
/clone\_lib="mth2"  
/note="Vector: pBelobAC11; Site 1: HindIII; Site 2: HindIII; Cook, D.R. and Kim, D.J, unpublished"

# ORIGIN

Alignment Scores:  
Pred. No.: 1.02e+03 Length: 485  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-10-014-101B-39 (1-7) x CG927858 (1-485)

QY 1 ValGlyGlyThrLeuSerAsn 7  
|||||

Db 397 GTTGGTGGCCTCTCTCTAAT 417  
|||||

# RESULT 18

BP031101

LOCUS

DEFINITION

BP031101 487 bp mRNA linear EST 19-AUG-2004  
BP031101 Lotus corniculatus var. japonicus flower Lotus corniculatus var. japonicus cDNA clone MF053g01\_f 3', mRNA sequence.

ACCESSION BP031101

VERSION BP031101.1 GI:45408261

KEYWORDS

SOURCE

ORGANISM

Lotus corniculatus var. japonicus (Lotus japonicus)  
Lotus corniculatus var. japonicus  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.

1 (bases 1 to 487)

Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.

Characteristics of the Lotus japonicus Gene Repertoire Deduced from Large-Scale Expressed Sequence Tag (EST) Analysis  
Plant Mol. Biol. 54 (3), 405-414 (2004)

Contact: Erika Asamizu

The First Laboratory for Plant Gene Research

Kazusa DNA Research Institute

Yana 1532-3, Kisarazu, Chiba 292-0812, Japan

Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

# FEATURES

source  
1. .487  
Location/Qualifiers  
/organism="Lotus corniculatus var. japonicus"  
/mol\_type="mRNA"  
/isolate="Miyakojima MG-20"  
/db\_xref="taxon:34305"  
/clone="MF053g01\_f"  
/tissue\_type="flower"

```

/clone_lib="Lotus corniculatus var. japonicus flower"
/clone_lib="Arabidopsis thaliana T-DNA insertion lines"
/ecotype="Col-0"
/notes="PCR was performed on DNA from Arabidopsis thaliana
plants (Ti) which were transformed with the T-DNA from
vector pAC106 (GenBank accession number: AJ537513). The
lines contain one or more T-DNA insertions. The DNA
fragment(s) resulting from the PCR were directly sequenced
to determine the genomic sequence flanking the insertion.
T-DNA derived sequences were removed."

ORIGIN
Alignment Scores:
Pred. No.: 1.03e+03 Length: 487
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-10-014-101B-39 (1-7) x BP031101 (1-487)

QY 1 ValGlyClyThrLeuSerAsn 7
Db 320 GTGGAGGAGCTCTTCCAAAT 340

RESULT 19
CR399680/c 491 bp DNA linear GSS 02-MAY-2004
LOCUS Arabidopsis thaliana T-DNA flanking sequence GK-820D08-025640,
DEFINITION genomic survey sequence.
ACCESSION CR399680
VERSION CR399680.1 GI:46940408
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Li, Y., Rosso, M.G., Strizhov, N., Viehoever, P. and Weisshaar, B.
TITLE GABI-Kat Simplesearch: a flanking sequence tag (FST) database for
the identification of T-DNA insertion mutants in Arabidopsis
thaliana
JOURNAL Bioinformatics 19 (11), 1441-1442 (2003)
MEDLINE 22755829
PUBMED 12874060
REFERENCE 2
AUTHORS Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K. and
Weisshaar, B.
TITLE An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for
flanking sequence tag-based reverse genetics
JOURNAL Plant Mol. Biol. 53 (1-2), 247-259 (2003)
MEDLINE 23117147
PUBMED 14756321
REFERENCE 3
AUTHORS Strizhov, N., Li, Y., Rosso, M.G., Viehoever, P., Dekker, K.A. and
Weisshaar, B.
TITLE High-throughput generation of sequence indexes from T-DNA
mutagenized Arabidopsis thaliana lines
JOURNAL Biotechniques 35 (6), 1164-1168 (2003)
PUBMED 14682050
REFERENCE 4
AUTHORS Strizhov, N., Rosso, M.G., Li, Y. and Weisshaar, B.
TITLE Direct Submission
JOURNAL Submitted (01-MAY-2004) Weisshaar B., Max-Planck-Institut fuer
Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT This sequence has been recovered from the left border of the T-DNA.
It indicates an insertion close to or within gene At2g19500.
Details on the protocols used for generation of the sequence are
described in References 1-3. The sequences are generated at the MPI
for Plant Breeding Research in the context of the GABI-Kat project.
GABI-Kat is part of the German Plant Genomics program designated
'GABI'. Information on line availability can be found at:
http://www.mpiz-koeln.mpg.de/GABI-Kat/.
Location/Qualifiers
1. .491
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="GK-820D08-025640"

FEATURES
source
Location/Qualifiers
1. .491
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0412P18"
/clone_lib="ZM 0.7-1.5_KB"
/notes="Vector: pBCSK; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

```

```

/clone_lib="Arabidopsis thaliana T-DNA insertion lines"
/ecotype="Col-0"
/notes="PCR was performed on DNA from Arabidopsis thaliana
plants (Ti) which were transformed with the T-DNA from
vector pAC106 (GenBank accession number: AJ537513). The
lines contain one or more T-DNA insertions. The DNA
fragment(s) resulting from the PCR were directly sequenced
to determine the genomic sequence flanking the insertion.
T-DNA derived sequences were removed."

ORIGIN
Alignment Scores:
Pred. No.: 1.04e+03 Length: 491
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-014-101B-39 (1-7) x CR399680 (1-491)

QY 1 ValGlyClyThrLeuSerAsn 7
Db 397 GTGGAGGAGCTTTCGAAT 377

RESULT 20
CR656142 504 bp DNA linear GSS 19-JUN-2003
LOCUS OGUDB93TH ZM 0.7-1.5_KB Zea mays genomic clone ZMMBMA0412P18,
DEFINITION genomic survey sequence.
ACCESSION CC656142
VERSION CC656142.1 GI:32059361
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1
AUTHORS Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other GSSs: OGUDB93TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1. .504
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0412P18"
/clone_lib="ZM 0.7-1.5_KB"
/notes="Vector: pBCSK; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

FEATURES
source
Location/Qualifiers
1. .504
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0412P18"
/clone_lib="ZM 0.7-1.5_KB"
/notes="Vector: pBCSK; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

```

US-10-014-101B-39 (1-7) x CC656142 (1-504)

QY 1 ValGlyGlyThrLeuSerAsn 7  
|||||  
Db 280 GTGGGGGAGCGTGTCCAAAT 300

RESULT 21

AQ827482/c  
LOCUS  
DEFINITION  
HS 5304 A1 C02 T7A RPCI-11 Human Male BAC Library Homo sapiens  
genomic clone Plate=980 Col=3 Row=E, genomic survey sequence.  
ACCESSION  
AQ827482  
VERSION  
AQ827482.1 GI:5793544  
KEYWORDS  
GSS.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 506)  
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,  
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and  
Hood,L.  
TITLE  
Sequence-tagged connectors: A sequence approach to mapping and  
scanning the human genome  
JOURNAL  
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
MEDLINE  
99380589  
PUBMED  
1049764  
COMMENT  
Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Clones are derived from the human BAC library RPCI-11. For BAC  
library availability, please contact Pieter de Jong  
(piederdejong.med.buffalo.edu). Clones may be purchased from  
BACPAC Resources ([http://bacpac.med.buffalo.edu/ordering\\_bac.htm](http://bacpac.med.buffalo.edu/ordering_bac.htm))  
or from Research Genetics (<http://www.htsc.washington.edu>)  
Plate: 880 row: E column: 3  
Seq primer: T7  
Class: BAC ends  
High quality sequence stop: 506.

FEATURES  
source

1..506  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="Plate=880 Col=3 Row=E"  
/sex="male"  
/notes="Vector: pBACE3.6; Site 1: EcoRI; Site 2: EcoRI;  
Male blood DNA was isolated from one randomly chosen donor  
and partially digested with a combination of EcoRI and  
EcoRI Methylase. Size selected DNA was cloned into the  
pBACE3.6 vector at EcoRI sites"

ORIGIN

Alignment Scores:  
Pred. No.: 1.07e+03 Length: 506  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0

US-10-014-101B-39 (1-7) x AQ827482 (1-506)

QY 1 ValGlyGlyThrLeuSerAsn 7  
|||||  
Db 367 GTAGGAGTACTCTCTCCAAAC 347

RESULT 22

CB640717/c  
LOCUS  
DEFINITION  
OSJNEal6A24.f OSJNEA Oryza sativa (japonica cultivar-group) cDNA  
clone OSJNEal6A24 5', mRNA sequence.  
ACCESSION  
CB640717  
VERSION  
CB640717.1 GI:296335708  
KEYWORDS  
EST.  
SOURCE  
Oryza sativa (japonica cultivar-group)  
ORGANISM  
Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Lillipsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaeae; Oryza.  
REFERENCE  
1 (bases 1 to 511)  
Jantassuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,  
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.  
Large-scale identification of ESTs involved in the interaction  
between rice and Magnaporthe grisea  
Unpublished (2003)  
COMMENT  
Contact: Rod Wing  
Arizona Genomics Institute  
University of Arizona  
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ  
85721-0088, USA  
Tel: 520 626 3967  
Fax: 520 621 9288  
Email: <http://genome.arizona.edu>  
PCR Primers  
FORWARD: atc agc ggc cgc gat cc  
BACKWARD: aat taa ccc tca cta aag gg  
Plate: 16 row: A column: 24  
Seq primer: atc agc ggc cgc gat cc.

FEATURES  
source

1..511  
Location/Qualifiers  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="OSJNEal6A24"  
/tissue type="Leaf"  
/dev stage="3 week"  
/lab\_host="DH108"  
/clone\_lib="OSJNEA"  
/note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:  
XhoI; 6 hrs after inoculation with Rice Blast (Che  
86061)"

ORIGIN

Alignment Scores:  
Pred. No.: 1.08e+03 Length: 511  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-014-101B-39 (1-7) x CB640717 (1-511)

QY 1 ValGlyGlyThrLeuSerAsn 7  
|||||  
Db 481 GTTGGAGGCACACTGTGGAAT 461

RESULT 23

BQ869165/c  
LOCUS  
DEFINITION  
BQ869165  
QGD5F19.yg.ab1 OG ABCDI lettuce salinas Lactuca sativa cDNA clone  
QGD5F19, mRNA sequence.  
ACCESSION  
BQ869165  
VERSION  
BQ869165.1 GI:22254922  
KEYWORDS  
EST.  
SOURCE  
Lactuca sativa  
ORGANISM  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Asterales; Asteraceae; Cichorioideae; Cichorieae; Lactuca.

1 (bases 1 to 515)

**REFERENCE**  
**AUTHORS**  
 Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L., Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J., Ellison, P., Kolkman, J., Slabaugh, M.S., Livingston, K., Zhou, Y., Lai, Z., Church, S., Jackson, L., and Bradford, K.

**TITLE**  
 Lettuce and Sunflower ESTs from the Compositae Genome Project

**JOURNAL**  
 http://compbio.ucdavis.edu/

**COMMENT**  
 Unpublished (2002)  
 Contact: Alexander Kozik [R.W.Michelmore]  
 Department of Vegetable Crops, R.W.Michelmore Lab  
 University of California at Davis (UCD)  
 Amundson Hall, UCD, Davis, CA 95616, USA  
 Tel: 1-(530)-742-1742  
 Fax: 1-(530)-752-9659  
 Email: akozik@ucdavis.edu [michelmore@veggmail.ucdavis.edu]  
 belongs to contig QG\_CA\_Contig6961, see http://cgdb.ucdavis.edu/ for details.

**FEATURES**  
 Plate: QGD5 row: F column: 19.

**source**  
 1. .515  
 Location/Qualifiers  
 /organism="Lactuca sativa"  
 /mol\_type="mRNA"  
 /cultivar="Salinas"  
 /db\_xref="taxon:4236"  
 /clone="QGD5P19"  
 /lab\_host="E.coli"  
 /clone\_lib="OG ABCDI lettuce salinas"  
 /note="Vector: pBRDNASfiAB; The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgdb.ucdavis.edu/ TAG TISSUE=flowers pre-fertilized  
 TAG LIB=OG ABCDI lettuce salinas  
 TAG\_SEQ=GGTTGACGGG"

**ORIGIN**

**Alignment Scores:**  
 Pred. No.: 1.09e+03 Length: 515  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 5 Gaps: 0

US-10-014-101b-39 (1-7) x BQ869165 (1-515)

**QY** 1 ValGlyGlyThrLeuSerAan 7  
 |||||  
**Db** 401 GTCGGCGGACGCTGTCTAAC 381

**RESULT 24**

**LOCUS** BX283649/c 518 bp mRNA linear EST 05-MAR-2003

**DEFINITION** BX283649 NIH MGC 21 Homo sapiens cDNA clone IMAGp958K03832 ; IMAGS:3956474, mRNA sequence.

**ACCESSION** BX283649

**VERSION** BX283649.1 GI:28848103

**KEYWORDS** EST.

**SOURCE** Homo sapiens (human)

**ORGANISM** Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

**REFERENCE** 1 (bases 1 to 518)

**AUTHORS** Ebert, L., Heil, O., Hennig, S., Neubert, P., Patsch, E., Peters, M., Radelof, U., Schneider, D. and Korn, B.

**TITLE**  
**JOURNAL**  
**COMMENT**

Human Unigeneset - RZPD3  
 Unpublished (2003)  
 Contact: Ina Rofls  
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
 Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany  
 RZPD: IMAGp958K03832  
 RZPDLIB: I.M.A.G.E. cDNA Clone Collection;  
 Human Unigeneset - RZPD3 (RZPDLIB No.972)  
 http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rofls  
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
 Heubnerweg 6, D-14059 Berlin, Germany  
 Tel: +49 30 32639 101  
 Fax: +49 30 32639 111  
 www.rzpd.de

This clone is available royalty-free from RZPD;  
 contact RZPD (clone@rzpd.de) for further information. Seq primer:  
 PCMV-M13u, Primer sequence: CGTTGTAAACGACGCCACT.

**FEATURES**

**source**  
 1. .518  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGp958K03832 ; IMAGE:3956474"  
 /tissue\_type="choriocarcinoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 21"  
 /note="Organ: placenta; Vector: pOTB7; Site: 1: XhoI; Site 2: BcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

**ORIGIN**

**Alignment Scores:**  
 Pred. No.: 1.1e+03 Length: 518  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 5 Gaps: 0

US-10-014-101b-39 (1-7) x BX283649 (1-518)

**QY** 1 ValGlyGlyThrLeuSerAan 7  
 |||||  
**Db** 85 GTGGCGGACGCTTCTTAAT 65

**RESULT 25**

**LOCUS** CR775437/c 520 bp mRNA linear EST 23-SEP-2004

**DEFINITION** DKFp469C0965\_r1 469 (synonym: pkidl) Pongo pygmaeus cDNA clone DKFp469C0965 5', mRNA sequence.

**ACCESSION** CR775437

**VERSION** CR775437.1 GI:52618686

**KEYWORDS** EST.

**SOURCE** Pongo pygmaeus (orangutan)

**ORGANISM** Pongo pygmaeus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.

**REFERENCE** 1 (bases 1 to 520)

**AUTHORS** Koehrer, K., Beyer, A., Mewes, H.W., Well, B., Amid, C., Oeanger, A., Pobo, G., Han, M. and Wiemann, S.

**TITLE****JOURNAL****COMMENT**

Unpublished (2004)

Contact: MIPS

MIPS

Ingolstaedter Landstr. 1, D-85764 Neuherberg, Germany  
 This is the 5' sequence of the clone insert. Clone from S. Wiemann,  
 Molecular Genome Analysis, German Cancer Research Center (DKFZ);

Email s.wiemann@kfz-heidelberg.de; mforschung GmbH in Berlin, Germany. Please contact RZPD for ordering:  
<http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp469C0965>  
 Further information about the clone and the sequencing project is available at <http://mips.gsf.de/projects/cdna/>.

#### FEATURES

source  
 location/Qualifiers  
 1..520  
 /organism="Pongo pygmaeus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9600"  
 /clone="DKFZp469C0965"  
 /dev\_stage="adult"  
 /tissue\_type="kidney"  
 /lab\_host="DH10B"  
 /clone\_lib="469 (synonym: pkid1)"  
 /note="Vector: pSport1\_Sfi; Site\_1: SfiI; Site\_2: SfiIb"

#### ORIGIN

Alignment Scores:  
 Pred. No.: 1..1e+03 Length: 520  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 7 Gaps: 0

US-10-014-101B-39 (1-7) x CR775437 (1-520)

Qy 1 ValGlyGlyThrLeuSerAsn 7

Db 129 GTGGCGGGAACGCTTCTTAAT 109

#### RESULT 26

BI498566 525 bp mRNA linear EST 08-JUL-2004  
 LOCUS sail9512.y1 Gm-c1053 Glycine max cDNA clone GENOME SYSTEMS CLONE  
 DEFINITION ID: Gm-c1053-3407 5' similar to TR:022213 O22213 T32G6.3 PROTEIN.  
 ; mRNA sequence.

ACCESSION BI498566

VERSION BI498566.1 GI:153337910

KEYWORDS EST.

SOURCE Glycine max (soybean)

ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

1 (bases 1 to 525)

REFERENCE  
 AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterson, R. and Wilson, R.

Public Soybean EST Project

Unpublished (1999)

CONTACT: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: [est@watson.wustl.edu](mailto:est@watson.wustl.edu)

When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. This clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone: 800 423 4163; email: [info@biogeneticservices.com](mailto:info@biogeneticservices.com))  
 High quality sequence stop: 391.

#### FEATURES

source  
 location/Qualifiers  
 1..525  
 /organism="Glycine max"  
 /mol\_type="mRNA"  
 /cultivar="Harosoy"

/db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1053-3407"  
 /tissue\_type="Whole seedling, 3 week old, greenhouse grown"  
 /lab\_host="DH10B"  
 /clone\_lib="Gm-c1053"

/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The Harosoy NIL was constructed and seed was provided by Dr. J. Specht, University of Nebraska (Shoemaker and Specht, 1995). The cDNA library was constructed from mRNA isolated from whole seedlings of 3 week old greenhouse grown plants. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site and a 3' anchor. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in cooperation with Dr. Paul Keim's laboratory at Northern Arizona University."

#### ORIGIN

Alignment Scores:  
 Pred. No.: 1..11e+03 Length: 525  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0

US-10-014-101B-39 (1-7) x BI498566 (1-525)

Qy 1 ValGlyGlyThrLeuSerAsn 7

Db 133 GTGGCGGGAACCTCTCGAAC 153

#### RESULT 27

AQ510474/c

LOCUS

DEFINITION

nbxb0095B17f CUGI Rice BAC Library Oryza sativa (japonica cultivar-group) genomic clone nbxb0095B17f, genomic survey sequence.

ACCESSION AQ510474

VERSION AQ510474.1 GI:4733078

KEYWORDS GSS.

SOURCE

ORGANISM

Oryza sativa (japonica cultivar-group)  
 Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzae; Oryza.

1 (bases 1 to 536)

REFERENCE

AUTHORS Wing, R.A. and Dean, R.A.

TITLE A BAC End Sequencing Framework to Sequence the Rice Genome

JOURNAL Unpublished (1998)

COMMENT

Contact: Wing RA

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288

Fax: 864 656 4293

Email: [rwing@clemson.edu](mailto:rwing@clemson.edu)

Seq primer: TAATACGACTCACTATAGGG

Class: BAC ends

High quality sequence stop: 346.

Location/Qualifiers

1..536

/organism="Oryza sativa (japonica cultivar-group)"

/mol\_type="genomic DNA"

/cultivar="Nipponbare"

/db\_xref="taxon:39947"

/clone="nbxb0095B17f"

/tissue\_type="Leaf"  
 /lab\_host="E. coli DH10B"  
 /clone\_lib="CUGI Rice BAC Library"  
 /notes="Vector: pBlueScript SK(-); Site 1: HindIII; Site 2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 Kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.14e+03 Length: 536  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservatative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 8 Gaps: 0

US-10-014-101B-39 (1-7) x AQ510474 (1-536)

QY 1 ValGlyGlyThrLeuSerAsn 7  
 Db 417 GTTGGAGGCACTGTGCTAAT 397

## RESULT 28

LOCUS AW224503 540 bp mRNA linear EST 18-MAY-2001  
 DEFINITION EST302946 tomato root, plants pre-anthesis, Cornell University  
 Lycopersicon esculentum cDNA clone CLE1A15 similar to Zea mays  
 cytochrome oxidase, mRNA sequence.

ACCESSION AW224503.1 GI:6536187  
 VERSION  
 KEYWORDS EST

SOURCE Lycopersicon esculentum (tomato)  
 ORGANISM

REFERENCE  
 AUTHORS van der Hoeven, R.S., Garvin, D., Matern, A.L., Holt, J.E., Liang, F., Upton, J., Hansen, I., Craven, M.B., Bowman, C.L., Ann, S., Ronning, C.M., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tankale, S.D.  
 TITLE Generation of ESTs from tomato root tissue  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: CUGI  
 Clemson University Genomics Institute  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 5 prime sequence.

FEATURES  
source

Location/Qualifiers  
 1..540  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="CLE1A15"  
 /tissue\_type="root"

/dev\_stage="plants in pre-anthesis stage"  
 /clone\_lib="tomato root, plants pre-anthesis, Cornell University"  
 /note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Tanksley; Tissue supplied by Dave Garvin (USDA-ARS, Ithaca, NY 14850)."

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.15e+03 Length: 540  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservatative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0

US-10-014-101B-39 (1-7) x AW224503 (1-540)

QY 1 ValGlyGlyThrLeuSerAsn 7  
 Db 180 GTGGAGGACTGTGCTAAT 200

## RESULT 29

CF603749

LOCUS BACCA01\_001055 Grape Berry pSPORT1 Library Vitis vinifera cDNA 5',  
 DEFINITION mRNA sequence.  
 ACCESSION CF603749  
 VERSION CF603749.1 GI:37184395

KEYWORDS EST  
 SOURCE Vitis vinifera  
 ORGANISM

REFERENCE  
 AUTHORS Moser, C., Segala, C., Fontana, P., Salakhudinov, I., Gatto, P., Pindo, M., Zyprian, E., Toepfer, R., Grando, M.S. and Velasco, R.  
 TITLE Expressed sequence tags from different organs of Vitis vinifera  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: Moser C

LABORATORIO DI GENETICA MOLECOLARE  
 ISTITUTO AGRARIO DI SAN MICHELE ALL'ADIGE (IASMA)  
 VIA E. MACH 1, SAN MICHELE ALL'ADIGE (TN), I-38010, ITALIA  
 TEL: 0039-0461-615314  
 FAX: 0039-0461-650956  
 EMAIL: claudio.moser@ismaa.it  
 The sequencing work has been funded by the 'Fondazione Cassa di  
 Risparmio di Trento e Rovereto'  
 High quality sequence stop: 542.

FEATURES  
source

Location/Qualifiers  
 1..542  
 /organism="Vitis vinifera"  
 /mol\_type="mRNA"  
 /cultivar="Pinot Noir"  
 /db\_xref="taxon:29760"  
 /sex="Hermaphrodite"  
 /dev\_stage="veraison"  
 /lab\_host="DH10B"  
 /clone\_lib="Grape Berry pSPORT1 Library"  
 /note="Organ: Berry; Vector: pSPORT1; Site\_1: NotI;  
 Site\_2: SalI"

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.15e+03 Length: 542  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservatative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 7 Gaps: 0

US-10-014-101B-39 (1-7) x CF603749 (1-542)

```

QY 1 ValGlyGlyThrLeuSerAsn 7
Db 412 GTGGGTGGACCCCTTCCGAAT 432

RESULT 30
LOCUS CG357974 551 bp DNA linear GSS 26-AUG-2003
DEFINITION OZBG58TV ZM_0.7-1.5 KB Zea mays genomic clone ZMMBMA0670119, genomic survey sequence.
ACCESSION CG357974
VERSION AL083412.1 GI:5284552
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 550)
AUTHORS Salanoubat,M., Choisne,N., Artiguenave,F., Brottier,P., Wincker,P.,
Samson,D., Saurin,W., Weissenbach,J. and Quetier,F.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 550)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (25-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

FEATURES
source
Location/Qualifiers
1..550
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="P6N21"
/clone_lib="IGF"
/notes="end : T7"

ORIGIN
Alignment Scores:
Pred. No.: 1.17e+03 Length: 550
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-014-101B-39 (1-7) x CNS000MU (1-550)

QY 1 ValGlyGlyThrLeuSerAsn 7
Db 377 GTGGGTGGACGTTATCAAC 397

RESULT 31
LOCUS CG357974 551 bp DNA linear GSS 26-AUG-2003
DEFINITION OZBG58TV ZM_0.7-1.5 KB Zea mays genomic clone ZMMBMA0670119, genomic survey sequence.
ACCESSION CG357974
VERSION CG357974.1 GI:34275241
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 551)
AUTHORS Whitlaw,C.A.; Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nurnberg,A., Robbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)

```

```

COMMENT
Other GSSs: OZBG58TH
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.

FEATURES
source
Location/Qualifiers
1..551
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0670119"
/clone_lib="ZM_0.7-1.5 KB"
/notes="Vector: pBCSK-; Site 1: HindII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Alignment Scores:
Pred. No.: 1.17e+03 Length: 551
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-014-101B-39 (1-7) x CG357974 (1-551)

QY 1 ValGlyGlyThrLeuSerAsn 7
Db 329 GTGGGCGGCACGCTCTCCAAT 349

RESULT 32
LOCUS CV057457/c 556 bp mRNA linear EST 24-AUG-2004
DEFINITION BNEL28a6 Barley EST endosperm library Hordeum vulgare subsp.
vulgare cDNA clone BNEL28a6 5' similar to putative cytokinin
dehydrogenase, mRNA sequence.
ACCESSION CV057457
VERSION CV057457.1 GI:51520596
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poaceae; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 556)
AUTHORS Ali,S, Holloway,B. and Taylor,W.C.
TITLE Normalisation of cereal endosperm EST libraries for structural and
functional genomic analysis
JOURNAL Plant Mol Biol. Rep. 18, 123-132 (2000)
COMMENT Contact: Bill Taylor
Commonwealth Scientific and Industrial Research Organisation
Division of Plant Industry.
CSIRO Plant Industry, GPO Box 1600, Canberra, ACT 2601, Australia
Tel: 61 2 6246 5223
Fax: 61 2 6246 5000
Email: Bill.taylor@csiro.au
Seq primer: M13 reverse primer
High quality sequence stop: 556.

FEATURES
source
Location/Qualifiers
1..556
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Himalaya"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="BNEL28a6"
/tissue_type="endosperm"
/dev_stage="developing endosperm tissue 10, 12, 15 dpa

```

(days post anthesis)"  
 /lab\_host="DH10B (Life Technology)"  
 /clone\_lib="Barley EST endosperm library"  
 /notes="Vector: Ziploxx; Site 1: Sal I; Site 2: Not I; mRNA was prepared from endosperm tissues of the barley cultivar Himalaya. cDNA was synthesized from pooled 10, 12, and 15 dpa endosperm using Not I-oligo(dT)18 primer/adaptor (Pharmacia Biotech), and then ligated to the Sal I-Not I site of Ziploxx vector (Life Technology) after adding a Sal I-Xho I adapter (Stratagene). Constructed by Shahjahan Ali and Bill Taylor."

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.18e+03 Length: 556  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 7 Gaps: 0

US-10-014-101B-39 (1-7) x CV057457 (1-556)

QY 1 ValGlyGlyThrLeuSerAan 7

DB 512 GTCGGCGGACGCTCTCCAAC 492

## RESULT 33

BI700658 559 bp mRNA linear EST 08-JUL-2004  
 LOCUS 8a127e03.y1 Gm-c1053 Glycine max cDNA clone GENOME SYSTEMS CLONE

DEFINITION ID: Gm-c1053-4445 5' similar to TR:022213 O22213 T32G6.3 PROTEIN.  
 // mRNA sequence.

ACCESSION BI700658

VERSION BI700658.1 GI:15663287

KEYWORDS EST.

SOURCE Glycine max (soybean)

ORGANISM Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE 1 (bases 1 to 559)

AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.  
 Public Soybean EST Project

UNPUBLISHED (1999)

TITLE Public Soybean EST Project

JOURNAL Public Soybean EST Project

COMMENT Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

When it has been determined, an EST from the other end of this clone is listed in the 'other ESTs on clone' field. This clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)  
 Seq primer: -40RP from Gibco

High quality sequence stop: 433.

FEATURES Location/Qualifiers

source

1..559

/organism="Glycine max"

/mol\_type="mRNA"

/cultivar="Harosoy"

/db\_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-c1053-4445"

/tissue\_type="Whole seedling, 3 week old, greenhouse

grown"

/lab\_host="DH10B"

/clone\_lib="Gm-c1053"

/notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The Harosoy NIL was constructed and seed was provided by Dr. J. Specht, University of Nebraska (Shoemaker and Specht, 1995). The cDNA library was constructed from mRNA isolated from whole seedlings of 3 week old greenhouse grown plants. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site and a 3' anchor. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in cooperation with Dr. Paul Keim's laboratory at Northern Arizona University."

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.19e+03 Length: 559  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0

US-10-014-101B-39 (1-7) x BI700658 (1-559)

QY 1 ValGlyGlyThrLeuSerAan 7

DB 166 GTCGGCGGACGCTCTCCAAT 186

## RESULT 34

BI736360

LOCUS

DEFINITION

AW736360 564 bp mRNA linear EST 07-SEP-2000

EST332279 KV3 Medicago truncatula cDNA clone pKV3-12E18, mRNA

sequence.

ACCESSION AW736360

VERSION AW736360.1 GI:7643143

KEYWORDS EST.

SOURCE Medicago truncatula (barrel medic)

ORGANISM Medicago truncatula  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

REFERENCE 1 (bases 1 to 564)

AUTHORS VandenBosch,K., Hurt,J., Moore,J., Beremand,P., Peng,H., Ellis,L., Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S., Holt,I.E. and Fraser,C.M.

ESTs from roots of Medicago truncatula after Rhizobium inoculation

UNPUBLISHED (1999)

TITLE Contact: VandenBosch K

JOURNAL Department of Plant Biology

COMMENT University of Minnesota

220 Biosci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA

Tel: 612 624 2755

Fax: 612 625 1738

Email: kvandenb@cbs.umn.edu

Texas A&M EST name: T255800e

TIGR sequence name: MTEAF33TK

More information is available at:

'http://chryslie.tamu.edu/medicago'

Seq primer: SKmod (CTA GAA CTA gtg gat CC).

FEATURES Location/Qualifiers

source

1..564

/organism="Medicago truncatula"

/mol\_type="mRNA"

/cultivar="genotype A17"

/db\_xref="taxon:3880"

/clone="pKV3-12E18"

/tissue\_type="Seedling roots"  
 /dev\_stage="3 days post-inoculation with Sinorhizobium  
 meliloti"  
 /lab\_host="E. coli strain XL0LR"  
 /clone\_lib="KV3"  
 /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:  
 XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA  
 was directionally ligated into the Unizap XR vector from  
 StrataGene and packaged using Gigapack III Gold packaging  
 extracts. Plasmids containing cDNA inserts were excised  
 from the recombinant lambda-Zap phage using Ex-assist  
 helper phage and propagated in XL0LR cells."

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.2e+03 Length: 564  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0

US-10-014-101B-39 (1-7) x AW736360 (1-564)

Qy 1 ValGlyGlyThrLeuSerAsn 7  
 Db 538 GTTGGTGGAACTCTCTCTAAT 558

## RESULT 35

CC022210/c

LOCUS CC022210 564 bp DNA linear GSS 01-APR-2003  
 DEFINITION 3591\_1\_28\_1\_G02.Y\_1\_3591 - RescueMu Grid P Zea mays genomic,  
 genomic survey sequence.

ACCESSION CC022210

VERSION CC022210.1 GI:29436283

KEYWORDS GSS

SOURCE Zea mays

ORGANISM Zea mays

REFERENCE 1 (bases 1 to 564)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoidae; Andropogoneae; Zea.

## AUTHORS

Walbot,V.

TITLE Maize genomic sequences found using engineered RescueMu transposon

JOURNAL Unpublished (2001)

COMMENT Contact: Walbot V

Department of Biological Sciences

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221

Email: walbot@stanford.edu

Plate: 3591.1.28.1 row: 15

Class: transposon-tagged

Location/Qualifiers

1..564

## FEATURES

source

/organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /cultivar="mixed background W23/A188/B73/K55"  
 /db\_xref="taxon:4577"  
 /tissue\_type="leaf"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"

/clone\_lib="3591 - RescueMu Grid P"  
 /note="Organ: leaf; Vector: RescueMu (engineered from  
 pBluescript backbone); Site 1: BamHI; Site 2: BglII;  
 RescueMu is a 4.9 kb, modified maize Mu transposon  
 designed to allow plasmid rescue from total genomic DNA.  
 Mu elements insert preferentially into transcription  
 units. For more information on RescueMu, go to the web  
 site 'www.zmndb.iastate.edu' and follow the links for  
 'RescueMu'. Grid P was grown at Molokai in 2002. DNA was  
 extracted from leaf strips, double digested using BamHI

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.2e+03 Length: 564  
 Score: 35.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 8 Gaps: 0

US-10-014-101B-39 (1-7) x CC022210 (1-564)

Qy 1 ValGlyGlyThrLeuSerAsn 7  
 Db 325 GTCGGTGGCACTTGTCTAAT 305

## RESULT 36

BU084470

LOCUS BU084470

DEFINITION BU084470

ACCESSION BU084470

VERSION BU084470.1

KEYWORDS GI:22525659

SOURCE EST.

ORGANISM Glycine max (soybean)

REFERENCE 1 (bases 1 to 566)

AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Corvett,V.,

Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,

Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,

Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,

Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,

McCann,R., Waterston,R. and Wilson,R.

Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

When it has been determined, an EST from the other end of this

clone is listed in the 'Other ESTs on clone' field. This clone is

available through: Biogenetic Services, 801 32nd Ave. Brookings, SD

57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)

Seq primer: -40RP from Gibco

High quality sequence stop: 421.

Location/Qualifiers

1..566

/organism="Glycine max"

/mol\_type="mRNA"

/cultivar="Clark"

/db\_xref="taxon:3847"

/clones="SOYBEAN CLONE ID: Gm-cl049-7252"

/tissue\_type="whole seedlings of greenhouse grown plants"

/dev\_stage="3 week old"

/lab\_host="DH10B"

/clone\_lib="Gm-cl049"

/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:

XhoI; The Clark NIL was constructed and seed was provided

by Dr. J. Specht, University of Nebraska (Shoemaker and

Specht, 1995). The cDNA library was constructed from mRNA

isolated from whole seedlings of 3 week old greenhouse

grown plants. Complementary DNA was synthesized from mRNA

and BglII, and ligated to form circular plasmids. DH10B  
 cells were transformed and then screened on LB plates with  
 ampicillin."

using a primer consisting of a poly(dT) sequence with a XhoI restriction site and a 3' anchor. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). The library was constructed in cooperation with Dr. Paul Keim's laboratory at Northern Arizona University."

## ORIGIN

Alignment Scores:  
Pred. No.: 1.21e+03 Length: 566  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 5 Gaps: 0

US-10-014-101B-39 (1-7) x BU084470 (1-566)

QY 1 ValGlyGlyThrLeuSerAsn 7  
|||||  
Db 133 GTGGAGGACCTTTCCAA 153

RESULT 37  
CL157422/c  
LOCUS  
DEFINITION  
104 345 10783550 114 31477 110 Sorghum methylation-filtered library (LibID: 104) Sorghum bicolor genomic clone 10783550, genomic survey sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Sorghum bicolor (sorghum)

REFERENCE  
AUTHORS  
Budiman, M.A., Flick, E., Jones, J., Nunberg, A., Citek, R.W., Robbins, D., Rohlfing, T., Bradford, K., Fries, J., McMenamy, J., Trani, L., Isak, A., Zimmerman, C., Lakey, N. and Bedell, J.A.,  
GeneThresher methylation filtered genomic sequences from Sorghum bicolor

JOURNAL  
COMMENT  
Unpublished (2004)

Contact: Bedell JA  
Orion Genomics, LLC  
4041 Forest Park Ave, St. Louis, MO 63108, USA  
Tel: 314 615 6979  
Fax: 314 615 5975

Email: jbedell@oriongenomics.com

Plate: 345 row: e column: 14

Seq primer: M13/pUC Forward

Class: shotgun

High quality sequence stop: 567.

## FEATURES

source  
1. 567  
Location/Qualifiers  
/organism="Sorghum bicolor"  
/mol\_type="genomic DNA"  
/cultivar="ATx623"  
/db\_xref="taxon:4558"  
/clone="10783550"  
/clone\_lib="Sorghum methylation-filtered library (LibID: 104)"  
/notes="Organ: leaf; Vector: pBCSK(-); Site 1: HincII, DNA prepared from purified nuclei was randomly sheared, end-repaired, size fractionated to enrich for the 0.5 to 5 kb fraction, ligated into HincII-digested pBCSK(-) vector and electroporated into E. coli cells. This is a methylation-filtered library."

## ORIGIN

Alignment Scores:  
Pred. No.: 1.21e+03 Length: 567  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-10-014-101B-39 (1-7) x CL157422 (1-567)

QY 1 ValGlyGlyThrLeuSerAsn 7  
|||||  
Db 451 GTGGGCGGACCTGTCTGAAC 431

RESULT 38

BZ706895/c

LOCUS

DEFINITION

BZ706895

OGBBD307C ZM\_0.7\_1.5\_KB Zea mays genomic clone ZMBMa0123B12, genomic survey sequence.

ACCESSION

BZ706895

VERSION

BZ706895.1

KEYWORDS

GSS.

SOURCE

Zea mays

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

AUTHORS

Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.

Consortium for Maize Genomics

Unpublished (2002)

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TF

Class: sheared ends.

Location/Qualifiers

1. 571

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone="ZMBMa0123B12"

/clone\_lib="ZM\_0.7\_1.5\_KB"

/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"

## ORIGIN

Alignment Scores:  
Pred. No.: 1.22e+03 Length: 571  
Score: 35.00 Matches: 7  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0

US-10-014-101B-39 (1-7) x BZ706895 (1-571)

QY 1 ValGlyGlyThrLeuSerAsn 7  
|||||  
Db 23 GTGGGCGGACCTGTCTAAT 3

RESULT 39

CO600263

LOCUS

DEFINITION

DG8-199k19 DG8-testis Canis familiaris cDNA 3', mRNA sequence.

ACCESSION

CO600263

VERSION

CO600263.1

GI:50445827

```

KEYWORDS EST.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE 1 (bases 1 to 573)
AUTHORS Schluter, T., Hermanns, J., Weindel, M., Schuette, D., Kranz, H.,
TITLE Dog arrayTAG cDNA clone collection
JOURNAL Unpublished (2004)
COMMENT Contact: Thomas Schluter
LION bioscience AG
Waldhoferstrasse 98, D-69123 Heidelberg, Germany
Tel: +49 6221 4038 150
Fax: +49 6221 4038 290
Email: Thomas.Schluter@lionbioscience.com.

FEATURES
source
1..573
/organism="Canis familiaris"
/mol_type="mRNA"
/strain="Beagle"
/db_xref="taxon:9615"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DRI0B"
/clone_lib="DG8-testis"
/note="Organ: testis; Vector: Dog pbluescript LION"

ORIGIN
Alignment Scores:
Pred. No.: 1.22e+03 Length: 574
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-10-014-101B-39 (1-7) x BP133582 (1-574)
QY 1 ValGlyGlyThrLeuSerAsn 7
Db 272 GTGGAGGACCTTGTGGAAT 252

Search completed: February 18, 2005, 07:00:50
Job time : 832.44 secs

/cultivar="Bright Yellow No.2"
/db_xref="taxon:4097"
/clone="BY5863"
/cell_line="BY-2"
/clone_lib="MAT001"
/note="Vector: pGEM-T easy; primer: M13 forward; mRNA
obtained from lag, log and stationary phase cells"

ORIGIN
Alignment Scores:
Pred. No.: 1.22e+03 Length: 574
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-10-014-101B-39 (1-7) x BP133582 (1-574)
QY 1 ValGlyGlyThrLeuSerAsn 7
Db 272 GTGGAGGACCTTGTGGAAT 252

Search completed: February 18, 2005, 07:00:50
Job time : 832.44 secs

KEYWORDS EST.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE 1 (bases 1 to 573)
AUTHORS Schluter, T., Hermanns, J., Weindel, M., Schuette, D., Kranz, H.,
TITLE Dog arrayTAG cDNA clone collection
JOURNAL Unpublished (2004)
COMMENT Contact: Thomas Schluter
LION bioscience AG
Waldhoferstrasse 98, D-69123 Heidelberg, Germany
Tel: +49 6221 4038 150
Fax: +49 6221 4038 290
Email: Thomas.Schluter@lionbioscience.com.

FEATURES
source
1..573
/organism="Canis familiaris"
/mol_type="mRNA"
/strain="Beagle"
/db_xref="taxon:9615"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DRI0B"
/clone_lib="DG8-testis"
/note="Organ: testis; Vector: Dog pbluescript LION"

ORIGIN
Alignment Scores:
Pred. No.: 1.22e+03 Length: 573
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 7 Gaps: 0

US-10-014-101B-39 (1-7) x CO600263 (1-573)
QY 1 ValGlyGlyThrLeuSerAsn 7
Db 359 GTGGAGGACGTTAAGTAAT 379

RESULT 40
BP133582/c
LOCUS BP133582 574 bp mRNA linear EST 16-JUL-2003
DEFINITION BP133582 MAT001 Nicotiana tabacum cDNA clone BY5863, mRNA sequence.
ACCESSION BP133582
VERSION BP133582.1 GI:32876467
KEYWORDS EST.
SOURCE Nicotiana tabacum (common tobacco)
ORGANISM Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Nicotiana.
1 (bases 1 to 574)
Matsuoka, K., Tashiro, G., Horiguchi, T., Demura, T. and Fukuda, H.
Profiling growth-phase dependent gene expression of tobacco BY-2
cells by comprehensive microarray analysis
Unpublished (2003)
JOURNAL
COMMENT Contact: Ken Matsuoka
Morphogenesis Research Group
RIKEN Plant Science Center
1-7-2 Suehirocho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9575
Fax: 81-45-503-9573
Email: by2@psc.riken.go.jp URL: http://mrg.psc.riken.go.jp/strc/
The cDNA library was constructed from mRNA isolated from lag (9 h),
log (72 h) and stationary (7 days) old BY-2 cells.
1..574
Location/Qualifiers
/organism="Nicotiana tabacum"
/mol_type="mRNA"

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 17, 2005, 21:59:48 ; Search time 1234.08 Seconds  
(without alignments)  
353.378 Million cell updates/sec

Title: US-10-014-101B-40  
Perfect score: 50  
Sequence: 1 VLGGLQCFC 9

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Zgapop 6.0 , Zgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2.1/USPTO.spool/US10014101/runat.16022005.075845.7936/app.query.fasta.1.796  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=40 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10014101 @CGN 1.1 2886 @runat.16022005.075845.7936 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.sm.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	50	100.0	5817	9 HSM807454	BX647310 Homo sapi
2	50	100.0	261203	2 AC126952	AC126952 Rattus no
3	48	96.0	196806	10 AC027184	AC027184 Mus muscu
4	48	96.0	197646	10 AC136376	AC136376 Mus muscu

5	48	96.0	255818	2 AC084055	AC084055 Mus muscu
C 6	46	92.0	155214	10 AL627406	AL627406 Mouse DNA
C 7	46	92.0	212540	2 AC102912	AC102912 Mus muscu
C 8	45	90.0	168177	2 AC023914	AC023914 Homo sapi
C 9	45	90.0	192682	10 AC107769	AC107769 Mus muscu
10	45	90.0	196902	9 AC105345	AC105345 Homo sapi
11	45	90.0	200000	2 AC007104	AC007104 Homo sapi
12	45	90.0	214284	9 AC116643	AC116643 Homo sapi
C 13	45	90.0	217225	10 AL513345	AL513345 Mouse DNA
C 14	45	90.0	236274	10 AC119853	AC119853 Mus muscu
C 15	45	90.0	277408	6 AC124136	AC124136 Mus muscu
C 16	44	88.0	1509	6 E07190	E07190 DNA encodin
17	44	88.0	68743	9 AL139349	AL139349 Human DNA
18	44	88.0	163410	9 AC006198	AC006198 Homo sapi
19	44	88.0	177725	2 AC105162	AC105162 Mus muscu
20	44	88.0	196529	10 AC102630	AC102630 Mus muscu
C 21	44	88.0	232605	1 AE017222	AE017222 Thermus t
C 22	44	88.0	279448	8 AY661656	AY661656 Sorghum b
C 23	43	86.0	5183	6 AR042369	AR042369 Sequence
C 24	43	86.0	5183	6 AR052277	AR052277 Sequence
C 25	43	86.0	5868	6 AR241188	AR241188 Sequence
C 26	43	86.0	5868	6 AR262645	AR262645 Sequence
27	43	86.0	6113	6 AX346724	AX346724 Sequence
C 28	43	86.0	6177	6 CO715456	CO715456 Sequence
C 29	43	86.0	6188	9 HUMHOX1	D45132 Homo sapien
30	43	86.0	6436	6 AX345583	AX345583 Sequence
C 31	43	86.0	6704	6 AX780015	AX780015 Sequence
32	43	86.0	7603	6 AX281287	AX281287 Sequence
33	43	86.0	7603	6 AX345210	AX345210 Sequence
C 34	43	86.0	7942	9 HSU17838	U17838 Homo sapien
35	43	86.0	9504	6 AX323793	AX323793 Sequence
C 36	43	86.0	55622	2 AC023042	AC023042 Homo sapi
37	43	86.0	59433	2 AC091556	AC091556 Homo sapi
38	43	86.0	76891	2 AC011925	AC011925 Homo sapi
39	43	86.0	110000	2 AC020850	Continuation (4 of
40	43	86.0	132446	2 AC024959	AC024959 Homo sapi
41	43	86.0	138054	2 AC146325	AC146325 Felis cat
42	43	86.0	151188	9 HS1177519	AL031277 Human DNA
C 43	43	86.0	165527	10 AL592403	AL592403 Mouse DNA
C 44	43	86.0	170336	10 AC125137	AC125137 Mus muscu
45	43	86.0	172793	2 AC023659	AC023659 Homo sapi

#### ALIGNMENTS

RESULT 1 HSM807454/c 5817 bp mRNA linear PRI 30-AUG-2003  
LOCUS HSM807454 Homo sapiens mRNA; cDNA DKFZp686D2193 (from clone DKFZp686D2193).  
DEFINITION BX647310  
ACCESSION BX647310  
VERSION BX647310.1 GI:34366338  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 5817)  
AUTHORS Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.  
CONSTRM The German Human cDNA Consortium  
TITLE Direct Submission  
JOURNAL Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY  
COMMENT Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.  
This clone (DKFZp686D2193) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further information about the clone and the sequencing project is available at http://mips.gsf.de/proj/cDNA/. Location/Qualifiers

```

source
1..5817
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686D2193"
/tissue_type="human retina"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_sfi; host
DH108; sites SfiIA + SfiIB"
/dev_stage="adult"
5781..5786
polyA_signal
5792
polyA_site
ORIGIN

Alignment Scores:
Pred. No.: 21.1 Length: 5817
Score: 50.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-014-101B-40 (1-9) x HSM807454 (1-5817)

Qy 1 ValLeuGlyGlyLeuGlyGlnPheCys 9
|||||
Db 1227 GTTCTGGTGCCGTCGGCCAGTTCGTC 1201

RESULT 2
AC126952
LOCUS
DEFINITION Rattus norvegicus clone CH230-1514, *** SEQUENCING IN PROGRESS ***,
8 unordered pieces.
AC126952
VERSION AC126952.5 GI:30522508
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 261203)
Muzny,D.,Marle,E.,Metzker,M.,Lee,S.,Abramson,S.,Adams,C.,Alder,J.,
Allen,C.,Allen,H.,Alsbrooks,S.,Amin,A.,Anguiano,D.,
Ayalebechi,V.,Ayagi,A.,Ayodeji,M.,Baca,E.,Baden,H.,
Baldwin,D.,Bandaranaike,D.,Barber,M.,Barnstead,M.,Benahmed,F.,
Biswalo,K.,Blair,J.,Blankenburg,K.,Blyth,P.,Brown,M.,
Bryant,N.,Buhay,C.,Burch,P.,Burrell,K.,Calderson,E.,
Cardenas,V.,Carter,K.,Cavazos,I.,Ceasar,H.,Center,A.,
Chacko,J.,Chavez,D.,Chen,G.,Chen,R.,Chen,Y.,Chen,Z.,Chu,J.,
Cleveland,C.,Cockrell,R.,Cox,C.,Coyle,M.,Cree,A.,D'Souza,L.,
Davila,M.L.,Davis,C.,Davy-Carroll,L.,De Anda,C.,Dederich,D.,
Deigado,O.,Denson,S.,Deramo,C.,Ding,Y.,Dinh,H.,Divya,K.,
Draper,H.,Dugan-Rocha,S.,Dunn,A.,Durbink,K.,Duval,B.,Eaves,K.,
Egan,A.,Escotto,M.,Eugene,C.,Evans,C.A.,Falls,T.,Fan,G.,
Fernandez,S.,Finley,M.,Flagg,N.,Forbes,L.,Poster,M.,Poster,P.,
Fraser,C.M.,Gabioli,A.,Garcia,R.,Garcia,A.,Garnier,T.,Garza,M.,
Gueorgeorgis,E.,Geer,K.,Gill,R.,Grady,M.,Guerra,W.,Guevara,W.,
Gunaratne,P.,Haaland,W.,Hamil,C.,Hamilton,C.,Hamilton,K.,
Harvey,Y.,Havlak,P.,Hawes,A.,Henderson,N.,Hernandez,J.,
Hernandez,R.,Hines,S.,Hladun,S.L.,Hodgson,A.,Hogues,M.,
Hollins,B.,Howells,S.,Hulvik,S.,Hume,J.,Idlebird,D.,Jackson,A.,
Jackson,L.,Jacob,L.,Jiang,H.,Johnson,B.,Johnson,R.,Jolivet,A.,
Karpachy,S.,Kelly,S.,Kelly,S.,Khan,Z.,King,L.,Kovar,C.,
Kowis,C.,Kraft,C.L.,Lebow,H.,Levan,J.,Lewis,L.,Li,Z.,Liu,J.,
Liu,J.,Liu,W.,Liu,Y.,London,P.,Longacre,S.,Lopez,J.,
Lorensuhewa,L.,Louisege,H.,Lozado,R.J.,Lu,X.,Ma,J.,
Maheshwari,M.,Mahindartne,M.,Mahmoud,M.,Malloy,K.,Mangum,A.,
Mangum,B.,Mapua,P.,Martin,K.,Martin,R.,Martinez,E.,
Mawhiney,S.,McLeod,M.P.,McNeill,T.Z.,Meenen,E.,
Milobavljivic,A.,Miner,G.,Minja,E.,Montemayor,J.,Moore,S.,
Morgan,M.,Morris,K.,Morris,S.,Munidas,M.,Murphy,M.,Nair,L.,
Nankervis,C.,Neal,D.,Newton,N.,Nguyen,N.,Norris,S.,
Nwaokeme,O.,Okwuonu,G.,Olarnpunaagoon,A.,Pal,S.,Parks,K.,
Pasternak,S.,Paul,H.,Perez,A.,Perez,L.,Pfannkoch,C.,
Plopper,F.,Poindexter,A.,Popovic,D.,Primus,E.,Pu,L.-L.,
Puzo,M.,Quiroz,J.,Rachlin,E.,Reeves,K.,Regier,M.A.,Reigh,R.,
Reilly,B.,Reilly,M.,Ren,Y.,Reuter,W.,Richards,S.,Riggs,F.,
Rives,C.,Rodkey,T.,Rojas,A.,Rose,M.,Rose,R.,Ruiz,S.J.,
Sanders,W.,Savery,G.,Scherer,S.,Scott,G.,Shatsman,S.,Shen,H.,
Shetty,J.,Shvartbeyn,A.,Sisson,I.,Sitter,C.D.,Smajs,D.,
Sneed,A.,Sodergren,E.,Song,X.-Z.,Sorelle,R.,Soosa,J.,
Steinle,M.,Strong,R.,Sutton,A.,Svatek,A.,Tabor,P.,Taylor,C.,
Taylor,T.,Thomas,N.,Thomas,D.,Tingey,A.,Trejos,Z.,Usmani,K.,
Valas,R.,Vera,V.,Villasana,D.,Waldron,L.,Walker,B.,Wang,J.,
Wang,Q.,Wang,S.,Warren,J.,Warren,R.,Wei,X.,White,F.,
Williams,G.,Willson,R.,Wlarczyk,R.,Wooden,H.,Worley,K.,
Wright,D.,Wright,R.,Wu,J.,Yakub,S.,Yen,J.,Yoon,L.,Yoon,V.,
Yu,F.,Zhang,J.,Zhou,X.,Zhou,X.,Zhao,S.,Dunn,D.,von
Niederhausern,A.,Weiss,R.,Smith,D.R.,Holt,R.A.,Smith,H.O.,
Weinstock,G.,and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 261203)
Worley,K.C.
Direct Submission
Submitted (11-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 261203)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 10, 2003 this sequence version replaced gi:25009166.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are oriented and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GQQA
Center clone name: CH230-1514
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 194451 bases at least Q40
Consensus quality: 199590 bases at least Q30
Consensus quality: 203294 bases at least Q20
Estimated insert size: 203856; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 250096: contig of 250096 bp in length
* 250097 250196: gap of unknown length
* 250197 251466: contig of 1270 bp in length

```

\* 251467 251566: gap of unknown length  
 \* 251567 252675: contig of 1109 bp in length  
 \* 252676 252775: gap of unknown length  
 \* 252776 253837: contig of 1062 bp in length  
 \* 253838 253937: gap of unknown length  
 \* 253938 255287: contig of 1350 bp in length  
 \* 255288 255387: gap of unknown length  
 \* 255388 257029: contig of 1642 bp in length  
 \* 257030 257129: gap of unknown length  
 \* 257130 259608: contig of 2479 bp in length  
 \* 259609 259708: gap of unknown length  
 \* 259709 261203: contig of 1495 bp in length.

## FEATURES

source

1. 261203  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 /clone="CH230-1514"  
 85376..86946  
 /notes="wgs\_contig"  
 139057..139631  
 /notes="clone boundary"  
 clone end:T7  
 site:EcoRI  
 end\_sequence:BH343764"

## ORIGIN

Alignment Scores:  
 Pred. No.: 706 Length: 261203  
 Score: 50.00 Matches: 9  
 Percent Similarity: 100.00% Conservat: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0

US-10-014-101B-40 (1-9) x AC126952 (1-261203)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
 |||||  
 Db 108136 GTTCTGGGTGGACTAGTCAATTGT 108162

## RESULT 3

AC027184 196806 bp DNA linear ROD 17-SEP-2004  
 LOCUS Mus musculus chromosome 14, clone RP23-138C14, complete sequence.  
 DEFINITION  
 AC027184  
 AC027184.15 GI:52219300  
 VERSION  
 HTG  
 KEYWORDS  
 SOURCE  
 Mus musculus (house mouse)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 196806)  
 Birren,B., Nusbaum,C. and Lander,E.  
 Mus musculus chromosome 14, clone RP23-138C14  
 Unpublished

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 2 (bases 1 to 196806)  
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
 Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,  
 Boguslavskiy,L., Bouckghalter,B., Brown,A., Burkett,G.,  
 Campopiano,A., Castle,A., Choquel,Y., Colangelo,M., Collins,S.,  
 Collymore,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S.,  
 Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,  
 Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,  
 Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,  
 Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,  
 Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehotsky,J.,  
 Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,  
 McCarthy,M., McEwan,P., McGovern,A., McKernan,K., McPheeters,R.,  
 Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J.,  
 Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
 O'Neil,D., Olivari,T.M., Oliver,J., Peterson,K., Pierre,N.,  
 Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,

Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,  
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
 Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,  
 Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,  
 Young,G., Zainoun,J., Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (28-MAR-2000) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 196806)  
 Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,  
 Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,  
 Bloom,T., Boguslavskiy,L., Bouckghalter,B., Camarata,J., Chang,J.,  
 Choquel,Y., Collymore,A., Cooke,P., Corum,B.,  
 DeArelano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,  
 Erickson,J., Faros,S., Ferreira,P., Fitzgerald,M., Gage,D.,  
 Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,  
 Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,  
 Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,  
 Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,  
 Maclean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,  
 McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,  
 Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,  
 O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,  
 Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,  
 Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,  
 Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,  
 Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,  
 Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,  
 Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,  
 Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (12-AUG-2004) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 4 (bases 1 to 196806)  
 Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,  
 Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,  
 Bloom,T., Boguslavskiy,L., Bouckghalter,B., Camarata,J., Chang,J.,  
 Choquel,Y., Collymore,A., Cooke,P., Corum,B.,  
 DeArelano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,  
 Erickson,J., Faros,S., Ferreira,P., Fitzgerald,M., Gage,D.,  
 Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,  
 Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,  
 Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,  
 Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,  
 Maclean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,  
 McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,  
 Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,  
 O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,  
 Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,  
 Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,  
 Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,  
 Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,  
 Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,  
 Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,  
 Zimmer,A. and Zody,M.

Direct Submission  
 Submitted (17-SEP-2004) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Sep 17, 2004 this sequence version replaced gi:51173695.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RN/RepeatMasker.html

----- Genome Center  
 Center: Whitehead Institute/MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@broad.mit.edu  
 ----- Project Information  
 Center project name: L7288  
 Center clone name: 138\_C\_14  
 ----- Location/Qualifiers  
 1. .196806

FEATURES  
source

```

/organism="Mus musculus"
/db_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="14"
/map="14"
/clone="RP23-138C14"
/clone_lib="RPCI-23 Female Mouse BAC"
complement(2884..2984)
/rpt_family="PB1D7"
3639..3730
/rpt_family="PB1D7"
3757..3943
/rpt_family="B2_Mm2"
complement(4007..5580)
/rpt_family="L1_MM"
5645..5732
/rpt_family="MMAR1"
5733..5878
/rpt_family="B1_MM"
5879..6717
/rpt_family="MMAR1"
6732..6871
/rpt_family="MMAR1"
complement(6847..6981)
/rpt_family="B1_MM"
7038..7175
/rpt_family="B1_MM"
complement(7963..8147)
/rpt_family="B3"
complement(8193..8278)
/rpt_family="PB1D7"
complement(8280..8397)
/rpt_family="B1F"
complement(8398..8423)
/rpt_family="ID_B1"
8450..8551
/rpt_family="(TG)n"
complement(8552..8697)
/rpt_family="B1_MM"
8698..8721
/rpt_family="(TG)n"
8722..8791
/rpt_family="(TG)n"
8933..8954
/rpt_family="AT_Rich"
complement(9117..9162)
/rpt_family="B4A"
complement(9163..9279)
/rpt_family="PB1D7"
10158..10226
/rpt_family="L2"
complement(10219..10305)
/rpt_family="L2"
11284..11388
/rpt_family="PB1D10"
12016..12406
/rpt_family="L2"
complement(12407..12553)
/rpt_family="B1_MM"
12554..12631
/rpt_family="L2"
complement(12919..13123)
/rpt_family="ORR1D"
complement(13124..13471)
/rpt_family="RLTR10"
complement(13472..13527)
/rpt_family="ORR1D"
13720..13914
/rpt_family="B3"
13930..14055
/rpt_family="B1_MM"
14386..14567
/rpt_family="ID_B1"

```

```

repeat_region 14878..14974
/rpt_family="L2"
complement(15018..15378)
/rpt_family="MTD"
15379..15591
/rpt_family="L2"
16811..16892
/rpt_family="PB1D9"
complement(16893..17288)
/rpt_family="MTC"
17641..17809
/rpt_family="B3"
17890..17907
/note="single clone coverage"
17893..17901
/note="30 qual single clone coverage"
17905..17936
/rpt_family="(CCCCCA)n"
17918..17924
/note="30 qual SNGL region"
complement(18176..18331)
/rpt_family="W12C"
complement(18310..18508)
/rpt_family="MT2A"

```

## Alignment Scores:

Pred. No.:	1.34e+03	Length:	196806
Score:	48.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	88.89%	Mismatches:	0
Query Match:	96.00%	Indels:	0
DB:	10	Gaps:	0

US-10-014-101b-40 (1-9) x AC027184 (1-196806)

QY 1 ValLeuGlyGlyLeuGlyClnPheCys 9

DB	23543	GTGATGGGTGGACTTGGTCAGTTCGT	23569
RESULT 4			

AC136376

LOCUS	AC136376	197646 bp	DNA	linear	ROD 30-JUL-2004
DEFINITION	Mus musculus chromosome 14 clone RP23-243M12, complete sequence.				
ACCESSION	AC136376				
VERSION	AC136376.3	GI:50839084			
KEYWORDS	HTG.				
SOURCE	Mus musculus (house mouse)				
ORGANISM					

REFERENCE 1 (bases 1 to 197646)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 Wilson,R.K.  
 TITLE The sequence of Mus musculus clone  
 JOURNAL Unpublished  
 AUTHORS 2 (bases 1 to 197646)  
 McPherson,J.D. and Waterston,R.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (31-OCT-2002) Genome Sequencing Center, 4444 Forest Park  
 Parkway, St. Louis, MO 63108, USA  
 REFERENCE 3 (bases 1 to 197646)  
 Wilson,R.K.  
 TITLE Direct Submission  
 JOURNAL Submitted (28-MAY-2004) Genome Sequencing Center, 4444 Forest Park  
 Parkway, St. Louis, MO 63108, USA  
 REFERENCE 4 (bases 1 to 197646)  
 Wilson,R.K.  
 TITLE Direct Submission  
 JOURNAL Submitted (30-JUL-2004) Genome Sequencing Center, 4444 Forest Park  
 Parkway, St. Louis, MO 63108, USA  
 COMMENT On Jul 30, 2004 this sequence version replaced gi:4777615.

## CENTER:

----- Genome Center -----  
 Center: Washington University Genome Sequencing Center

Center code: WUGSC  
Web site: <http://genome.wustl.edu>  
Contact: [submissions@wustl.edu](mailto:submissions@wustl.edu)  
----- Project Information -----  
Center project name: M.BA0243M12  
----- Location/Qualifiers -----  
FEATURES  
source  
1. 197646  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="14"  
/clone="RP23-243M12"

ORIGIN  
Alignment Scores:  
Pred. No.: 1.34e+03 Length: 197646  
Score: 48.00 Matches: 8  
Percent Similarity: 100.00% Conservatives: 1  
Best Local Similarity: 88.89% Mismatches: 0  
Query Match: 96.00% Indels: 0  
DB: 10 Gaps: 0

US-10-014-101b-40 (1-9) x AC136376 (1-197646)  
Qy 1 ValLeuGlyLeuGlyGlnPheCys 9  
||||:|||||  
Db 173706 GTCATGGTGGAGCTTGTGTCAGTCTGT 173732

RESULT 5  
AC084055  
LOCUS  
DEFINITION AC084055 255818 bp DNA linear HTG 08-JUN-2002  
Mus musculus chromosome 2 clone RP23-359C19 strain C57BL6/J,  
WORKING DRAFT SEQUENCE, 37 unordered pieces.  
AC084055  
AC084055.6 GI:21358695  
VERSION  
HTG; HTGS PHASE1; HTGS DRAFT.  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
REFERENCE  
1 (bases 1 to 255818)  
AUTHORS Montgomerly, K.T., Grills, G., Han, J., Chiu, D., Decker, J., Fusina, M.,  
Goltz, J., Haider, A., Hall, L., Ioshikhes, I.P., Lee, E., Long, J.,  
Perera, A.M., Shim, C., Thomas, E.C. and Kucherlapati, R.  
TITLE High Throughput Mouse Sequencing  
JOURNAL Unpublished  
REFERENCE  
2 (bases 1 to 255818)  
AUTHORS Montgomerly, K.T., Grills, G., Han, J., Chiu, D., Decker, J., Fusina, M.,  
Goltz, J., Haider, A., Hall, L., Ioshikhes, I.P., Lee, E., Long, J.,  
Perera, A., Shim, C., Thomas, E. and Kucherlapati, R.  
TITLE Direct Submission  
JOURNAL Submitted (12-OCT-2000) Department of Molecular Genetics, Albert  
Einstein College of Medicine Genome Center, 1300 Morris Park Ave.,  
Bronx, NY 10461, USA  
REFERENCE  
3 (bases 1 to 255818)  
AUTHORS Montgomerly, K.T., Grills, G., Han, J., Chiu, D., Decker, J., Fusina, M.,  
Goltz, J., Haider, A., Hall, L., Ioshikhes, I.P., Lee, E., Long, J.,  
Perera, A., Shim, C., Thomas, E. and Kucherlapati, R.  
TITLE Direct Submission  
JOURNAL Submitted (08-JUN-2002) Harvard Partners Center for Genetics and  
Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA  
02139, USA  
COMMENT  
On Jun 8, 2002 this sequence version replaced gi:213066620.  
-----Genome Center  
Center: Harvard Partners Genome Center  
Center Code: HPGC  
Web site: <http://www.hpcgc.org/sequence/mouse.html>  
Contact: [hpcgc@mendel.mgh.harvard.edu](mailto:hpcgc@mendel.mgh.harvard.edu)  
-----Summary Statistics-----  
Center project name: AEK  
Sequencing vector: pUC18; L08752  
Chemistry: Dye-terminator Big Dye; 100\*

\*Consensus quality: 238808 at least Q20  
\*Consensus quality: 233351 at least Q30  
\*Consensus quality: 226354 at least Q40  
\*\*Estimated insert size: agarose-FP - N/A  
\*\*Estimated insert size: 255098 - sum-of-contigs  
Quality coverage: agarose-FP - N/A  
Quality coverage: 6.8 x in Q20 bases; sum-of-contigs estimation  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
consists of 37 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
be preserved.  
\* 1 69107: contig of 69107 bp in length  
\* 69108 69127: gap of unknown length  
\* 69128 122703: contig of 53576 bp in length  
\* 122704 122723: gap of unknown length  
\* 122724 149769: contig of 27046 bp in length  
\* 149770 149789: gap of unknown length  
\* 149790 165917: contig of 16128 bp in length  
\* 165918 165937: gap of unknown length  
\* 165938 177120: contig of 11183 bp in length  
\* 177121 177140: gap of unknown length  
\* 177141 186790: contig of 9650 bp in length  
\* 186791 186810: gap of unknown length  
\* 186811 195013: contig of 8203 bp in length  
\* 195014 195033: gap of unknown length  
\* 195034 201988: contig of 6955 bp in length  
\* 201989 202008: gap of unknown length  
\* 202009 206594: contig of 4586 bp in length  
\* 206595 206614: gap of unknown length  
\* 206615 209494: contig of 2880 bp in length  
\* 209495 209514: gap of unknown length  
\* 209515 210396: contig of 881 bp in length  
\* 210397 210415: gap of unknown length  
\* 210416 211817: contig of 1401 bp in length  
\* 211818 211836: gap of unknown length  
\* 211837 213977: contig of 2141 bp in length  
\* 213978 213997: gap of unknown length  
\* 213998 216349: contig of 2352 bp in length  
\* 216350 216369: gap of unknown length  
\* 216370 218294: contig of 1925 bp in length  
\* 218295 218314: gap of unknown length  
\* 218315 220170: contig of 1856 bp in length  
\* 220171 220190: gap of unknown length  
\* 220191 223017: contig of 2827 bp in length  
\* 223018 223037: gap of unknown length  
\* 223038 226049: contig of 3012 bp in length  
\* 226050 226069: gap of unknown length  
\* 226070 227531: contig of 1462 bp in length  
\* 227532 227551: gap of unknown length  
\* 227552 230223: contig of 2672 bp in length  
\* 230224 230243: gap of unknown length  
\* 230244 232298: contig of 2055 bp in length  
\* 232299 232318: gap of unknown length  
\* 232319 233599: contig of 1281 bp in length  
\* 233600 233619: gap of unknown length  
\* 233620 236005: contig of 2386 bp in length  
\* 236006 236025: gap of unknown length  
\* 236026 236902: contig of 877 bp in length  
\* 236903 236922: gap of unknown length  
\* 236923 238521: contig of 1929 bp in length  
\* 238522 238871: gap of unknown length  
\* 238872 240684: contig of 1813 bp in length  
\* 240685 240704: gap of unknown length  
\* 240705 242179: contig of 1475 bp in length  
\* 242180 242199: gap of unknown length  
\* 242200 243272: contig of 1073 bp in length  
\* 243273 243292: gap of unknown length  
\* 243293 244114: contig of 822 bp in length

```

* 244115 244134: gap of unknown length
* 244135 246257: contig of 2123 bp in length
* 246258 246277: gap of unknown length
* 246278 248050: contig of 1773 bp in length
* 248051 248070: gap of unknown length
* 248071 248889: contig of 819 bp in length
* 248890 248909: gap of unknown length
* 248910 250641: contig of 1732 bp in length
* 250642 250661: gap of unknown length
* 250662 252029: contig of 1368 bp in length
* 252030 252049: gap of unknown length
* 252050 253649: contig of 1600 bp in length
* 253650 253669: gap of unknown length
* 253670 254711: contig of 1042 bp in length
* 254712 254731: gap of unknown length
* 254732 255818: contig of 1087 bp in length.
FEATURES
    source
        Location/Qualifiers
            1..255818
                /organism="Mus musculus"
                /mol_type="genomic DNA"
                /strain="C57BL6/J"
                /db_xref="taxon:10090"
                /chromosome="2"
                /clone="RP23-359C19"
                /sex="male"
            1..69107
                /note="assembly_name:Contig124"
            69128..122703
                /note="assembly_name:Contig123"
            122724..149769
                /note="assembly_name:Contig122"
                clone_end:SP6
                vector_side:right
            149790..165917
                /note="assembly_name:Contig121"
            165938..177120
                /note="assembly_name:Contig120"
            177141..186790
                /note="assembly_name:Contig119"
            186811..195013
                /note="assembly_name:Contig118"
            195034..201988
                /note="assembly_name:Contig117"
            202009..206594
                /note="assembly_name:Contig116"
            206615..209494
                /note="assembly_name:Contig115"
            209515..210395
                /note="assembly_name:Contig114"
            210416..211816
                /note="assembly_name:Contig113"
            211837..213977
                /note="assembly_name:Contig112"
            213998..216349
                /note="assembly_name:Contig111"
            216370..218294
                /note="assembly_name:Contig110"
            218315..220170
                /note="assembly_name:Contig109"
            220191..223017
                /note="assembly_name:Contig108"
            223038..226049
                /note="assembly_name:Contig107"
            226070..227531
                /note="assembly_name:Contig106"
            227552..230223
                /note="assembly_name:Contig105"
            230244..232298
                /note="assembly_name:Contig104"
            232319..233599
                /note="assembly_name:Contig103"
            233620..236005
                /note="assembly_name:Contig102"

```

```

misc_feature 236026..236902
    /note="assembly_name:Contig101"
misc_feature 236923..238851
    /note="assembly_name:Contig100"
misc_feature 238872..240684
    /note="assembly_name:Contig99"
misc_feature 240705..242179
    /note="assembly_name:Contig98"
misc_feature 242200..243272
    /note="assembly_name:Contig97"
misc_feature 243293..244114
    /note="assembly_name:Contig96"

Alignment Scores:
Pred. No.:      1.7e+03      Length:      255818
Score:          48.00        Matches:      8
Percent Similarity: 100.00%   Conservative: 1
Best Local Similarity: 88.89%   Mismatches:  0
Query Match:     96.00%       Indels:       0
DB:              2           Gaps:         0

```

US-10-014-101B-40 (1-9) x AC084055 (1-255818)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9

Db 76857 GTCATGGGTGGACTGGTCAGTTCTGT 76883

#### RESULT 6

AL627406/c

LOCUS AL627406 155214 bp DNA linear ROD 09-JUL-2002  
DEFINITION Mouse DNA sequence from clone RP23-188A12 on chromosome 4, complete sequence.

ACCESSION AL627406

VERSION AL627406.15 GI:21727347

KEYWORDS HTG.

SOURCE Mus musculus (house mouse)

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 155214)

Tracey, A.

Direct Submission

Submitted (09-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

On Jul 10, 2002 this sequence version replaced gi.21711854.

#### COMMENT

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)

-----

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclones; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at

[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) RP23-188A12 is

from the RPI-23 Mouse PAC Library

constructed by the group of Pieter de Jong.

For further details see <http://www.chori.org/bacpac/home.htm>

FEATURES  
 Source  
 VECTOR: pBACe3.6  
 Location/Qualifiers  
 1. 155214  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="4"  
 /clone="RP23-188A12"  
 /clone\_lib="RPCI-23"

ORIGIN  
 Alignment Scores: 2.64e+03 Length: 155214  
 Score: 46.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 88.89% Mismatches: 0  
 Query Match: 92.00% Indels: 0  
 DB: 10 Gaps: 0

US-10-014-101B-40 (1-9) x AL627406 (1-155214)

QY 1 ValLeuGlyLeuGlyGlnPheCys 9  
 |||||  
 Db 33671 GTTTGGGTGGTCTGGGTAGTTCTGT 33645

RESULT 7  
 AC102912/c  
 LOCUS  
 DEFINITION Mus musculus clone RP24-463E18, WORKING DRAFT SEQUENCE, 5 unordered pieces.  
 AC102912  
 HTG: HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
 Mus musculus (house mouse)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 212540)  
 Birren, B., Nussbaum, C. and Lander, E.  
 Mus musculus, clone RP24-463E18  
 Unpublished  
 2 (bases 1 to 212540)  
 Birren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Collangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
 Direct Submission  
 Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 212540)  
 Birren, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArelano, K., Collamore, A., Cook, A., Dooly, K., Dorris, L., Erickson, J., Faro, S., Diaz, J.S., Dodge, S., Dooly, K., Dorris, L., Erickson, J., Faro, S.,

TITLE  
 JOURNAL  
 COMMENT  
 On Mar 2, 2003 this sequence version replaced gi:2381871.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html  
 ----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 ----- Project Information  
 Center project name: L20006  
 Center clone name: 463\_E\_18  
 ----- Summary Statistics  
 Sequencing vector: Plasmid; n/a; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 211193 bases at least Q40  
 Consensus quality: 211786 bases at least Q30  
 Consensus quality: 211996 bases at least Q20  
 Insert size: 178000; agarose-fp  
 Insert size: 212140; sum-of-contigs  
 Quality coverage: 8.9 in Q20 bases; agarose-fp  
 Quality coverage: 7.4 in Q20 bases; sum-of-contigs  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 5 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1 21291: contig of 21291 bp in length  
 \* 21292 21391: gap of 100 bp  
 \* 21392 41525: contig of 20134 bp in length  
 \* 41526 41626: gap of 100 bp  
 \* 41626 83953: contig of 42328 bp in length  
 \* 83954 84053: gap of 100 bp  
 \* 84054 128582: contig of 44529 bp in length  
 \* 128583 128683: gap of 100 bp  
 \* 128683 212540: contig of 83858 bp in length.  
 FEATURES  
 source  
 1. 212540  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /clone="RP24-463E18"  
 /clone\_lib="RPCI-24 Male Mouse BAC"  
 misc\_feature  
 1. 21291  
 /note="assembly\_fragment"  
 misc\_feature  
 21392..41525  
 /note="assembly\_fragment"  
 misc\_feature  
 41626..83953  
 /note="assembly\_fragment"  
 misc\_feature  
 84054..128582

```

/note="assembly_fragment"
128683..212540
/note="assembly_fragment"

misc_feature
ORIGIN

Alignment Scores:
Pred. No.: 3.52e+03 Length: 212540
Score: 46.00 Matches: 8
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 88.89% Mismatches: 0
Query Match: 92.00% Indels: 0
DB: 2 Gaps: 0

US-10-014-101b-40 (1-9) x AC102912 (1-212540)
Qy 1 ValLeuGlyLeuGlyGlnPheCys 9
Db 122307 GTTTGGTGTCTGGGTAGTCTGT 122281

RESULT 8
AC023914/c
LOCUS AC023914 168177 bp DNA linear HTG 20-SEP-2000
DEFINITION Homo sapiens clone RP11-44K20, WORKING DRAFT SEQUENCE, 22 unordered
pieces.
AC023914
AC023914.3 GI:10198362
VERSION HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 168177)
Birren,B., Linton,L., Nussbaum,C. and Lander,E.
Homo sapiens, clone RP11-44K20
Unpublished
2 (bases 1 to 168177)
Birren,B., Linton,L., Nussbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beda,F., Boguslavsky,L.,
Boukagalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
Choepe,Y., Collangelo,M., Collins,S., Collimore,A., Cooke,P.,
DeArellano,K., Dewar,K., Dodge,S., Domino,M., Doyle,M.,
Fenster,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Landers,T., Largocque,K., Lehoczy,J., Levine,R.,
Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M.,
McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Olivari,T.M.,
Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C.,
Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B.,
Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and
Zody,M.
Direct Submission
Submitted (20-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 20, 2000 this sequence version replaced gi:9102875.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L1087
Center clone name: 44_K_20
----- Summary Statistics
FEATURES
source
1..168177
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="RP11-44K20"
/clone_lib="RP11-11 Human Male BAC"
1..1263
/note="assembly_fragment"
clone end:SP6
vector_side:left"

```

Sequencing vector: M13; M77815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 157528 bases at least Q40  
Consensus quality: 162841 bases at least Q30  
Consensus quality: 164870 bases at least Q20  
Insert size: 176000; agarose-fp  
Insert size: 166077; sum-of-contigs  
Quality coverage: 4.3 in Q20 bases; agarose-fp  
Quality coverage: 4.5 in Q20 bases; sum-of-contigs

-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 22 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1263: contig of 1263 bp in length  
1264 1363: gap of 100 bp  
1364 2399: contig of 1036 bp in length  
2400 2499: gap of 100 bp  
2500 3713: contig of 1214 bp in length  
3714 3813: gap of 100 bp  
3814 27125: contig of 23312 bp in length  
27126 27225: gap of 100 bp  
27226 28854: contig of 1629 bp in length  
28855 28954: gap of 100 bp  
28955 31625: contig of 2671 bp in length  
31626 31725: gap of 100 bp  
31726 34084: contig of 2359 bp in length  
34085 34184: gap of 100 bp  
34185 36582: contig of 2398 bp in length  
36583 36682: gap of 100 bp  
36683 39629: contig of 2947 bp in length  
39630 39729: gap of 100 bp  
39730 43900: contig of 4171 bp in length  
43901 44000: gap of 100 bp  
44001 46837: contig of 2837 bp in length  
46838 46937: gap of 100 bp  
46938 52599: contig of 5662 bp in length  
52600 52699: gap of 100 bp  
52700 60335: contig of 7536 bp in length  
60336 60336: contig of 100 bp  
60336 67600: contig of 7265 bp in length  
67601 67700: gap of 100 bp  
67701 76219: contig of 8519 bp in length  
76220 76319: gap of 100 bp  
76320 85003: contig of 8684 bp in length  
85004 85103: gap of 100 bp  
85104 95875: contig of 10772 bp in length  
95876 95876: gap of 100 bp  
95876 108080: contig of 12105 bp in length  
108081 108180: gap of 100 bp  
108181 126996: contig of 18816 bp in length  
126997 127096: gap of 100 bp  
127097 145024: contig of 17928 bp in length  
145025 145124: gap of 100 bp  
145125 166554: contig of 21430 bp in length  
166555 166555: gap of 100 bp  
166555 168177: contig of 1523 bp in length.

Location/Qualifiers  
1..168177  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="RP11-44K20"  
/clone\_lib="RP11-11 Human Male BAC"  
1..1263  
/note="assembly\_fragment"  
clone end:SP6  
vector\_side:left"

```

misc_feature 1364..2399
/notes="assembly_fragment"
misc_feature 2500..3713
/notes="assembly_fragment"
misc_feature 3814..27125
/notes="assembly_fragment"
misc_feature 27226..28854
/notes="assembly_fragment"
misc_feature 28955..31625
/notes="assembly_fragment"
misc_feature 31726..34084
/notes="assembly_fragment"
misc_feature 34185..36582
/notes="assembly_fragment"
misc_feature 36683..39629
/notes="assembly_fragment"
misc_feature 39730..43900
/notes="assembly_fragment"
misc_feature 44001..46837
/notes="assembly_fragment"
misc_feature 46938..52599
/notes="assembly_fragment"
misc_feature 52700..60235
/notes="assembly_fragment"
misc_feature 60336..67600
/notes="assembly_fragment"
misc_feature 67701..76219
/notes="assembly_fragment"
misc_feature 76320..85003
/notes="assembly_fragment"
misc_feature 85104..95875
/notes="assembly_fragment"
misc_feature 95976..108080
/notes="assembly_fragment"
misc_feature 108181..126996
/notes="assembly_fragment"
misc_feature 127097..145024
/notes="assembly_fragment"
misc_feature 145125..166554
/notes="assembly_fragment"
misc_feature 166655..168177
/notes="assembly_fragment"
clone_end:T7
vector_side:right"

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 4,45e+03 Length: 168177
Score: 45.00 Matches: 8
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 1
Query Match: 90.00% Indels: 0
DB: 2 Gaps: 0

```

US-10-014-101B-40 (1-9) x AC023914 (1-168177)

QY 1 ValLeuGlyLeuGlyGlnPheCys 9  
 |||||  
 Db 24360 GTGTGGGGGGCTTGGTAGCTTCTGC 24334

## RESULT 9

AC107769/c AC107769 192682 bp DNA linear ROD 30-SEP-2003  
 LOCUS Mus musculus chromosome 16, clone RP23-130N20, complete sequence.

DEFINITION AC107769

ACCESSION AC107769

VERSION AC107769.7 GI:33342391

KEYWORDS HTG.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 192682)

AUTHORS Birren,B., Nusbaum,C. and Lander,E.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Mus musculus chromosome 16, clone RP23-130N20

Unpublished

2 (bases 1 to 192682)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,  
 Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,  
 Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,  
 Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,  
 Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,  
 Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,  
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
 Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
 Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,  
 Landers,T., Lehotzky,J., Levine,R., Liu,G., Maclean,C.,  
 Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,  
 McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,  
 Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,  
 Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,  
 Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,  
 Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,  
 Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,  
 Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
 Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,  
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,  
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 192682)

REFERENCE  
AUTHORS

Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,  
 Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,  
 Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,  
 Collymore,A., Cook,A., Cooke,P., Corum,B., DeArelano,K.,  
 Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,  
 Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,  
 Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,  
 Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
 Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,  
 Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., Maclean,C.,  
 Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,  
 Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,  
 Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,  
 O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,  
 Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,  
 Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,  
 Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,  
 Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M.,  
 Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,  
 Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (04-JUL-2003) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 192682)

REFERENCE  
AUTHORS

Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,  
 Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,  
 Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,  
 Collymore,A., Cook,A., Cooke,P., Corum,B., DeArelano,K.,  
 Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,  
 Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,  
 Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,  
 Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
 Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,  
 Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., Maclean,C.,  
 Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,  
 Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,  
 Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,  
 O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,  
 Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,  
 Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,  
 Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,  
 Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M.,  
 Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,  
 Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

**TITLE**  
**JOURNAL**  
**REFERENCE**  
**AUTHORS**

Direct Submission  
 Submitted (30-JUL-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 5 (bases 1 to 192682)  
 Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kellis, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, X., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meidrim, J., Meneu, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

**TITLE**  
**JOURNAL**  
**COMMENT**

Direct Submission  
 Submitted (30-SEP-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Jul 30, 2003 this sequence version replaced gi:32451548.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: <http://www-seq.wi.mit.edu>  
 Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)  
 ----- Project Information  
 Center project name: L20430  
 Center clone name: 130\_N\_20

**FEATURES**  
 source

Location/Qualifiers  
 1..192682  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="16"  
 /map="16"  
 /clone="RP23-130N20"  
 /clone\_lib="RPCI-23 Female Mouse BAC"

**misc\_feature**  
 1..4987  
 /note="wgs end extension"

clone\_end:SP6"

87..125  
 /rpt\_family="CT-rich"

complement(126..266)  
 /rpt\_family="B1f"

complement(152..271)  
 /rpt\_family="ID\_B1"

complement(373..509)  
 /rpt\_family="B1\_MM"

675..1079  
 /rpt\_family="RLTR9B"

681..729  
 /note="single clone coverage"

1085..1500  
 /rpt\_family="ETnERV"

1483..2442  
 /rpt\_family="ETnERV"

2468..4040  
 /rpt\_family="ETnERV"

4009..4494  
 /rpt\_family="RLTR9B-int"

complement(4489..4750)

**repeat\_region**  
 complement(4495..4684)  
 /rpt\_family="B2\_MM"

**repeat\_region**  
 4685..5657  
 /rpt\_family="RLTR9B-int"

**misc\_feature**  
 4988..4992  
 /note="clone boundary"

clone\_end:SP6  
 site:ECORI"

**repeat\_region**  
 5868..5941  
 /rpt\_family="RLTR9B-int"

**repeat\_region**  
 5942..6340  
 /rpt\_family="RLTR9B"

**repeat\_region**  
 6709..6796  
 /rpt\_family="MIR"

**repeat\_region**  
 complement(6917..7021)  
 /rpt\_family="PB1D10"

**repeat\_region**  
 7044..7091  
 /rpt\_family="(TG)n"

**repeat\_region**  
 complement(7204..7352)  
 /rpt\_family="B1\_MM"

**repeat\_region**  
 complement(7358..7556)  
 /rpt\_family="B3A"

**repeat\_region**  
 complement(7557..7760)  
 /rpt\_family="B4A"

**repeat\_region**  
 7767..7822  
 /rpt\_family="AT-rich"

**repeat\_region**  
 7840..7915  
 /rpt\_family="ID4"

**repeat\_region**  
 8020..8167  
 /rpt\_family="B1\_MM"

**repeat\_region**  
 8723..8874  
 /rpt\_family="RSINE1"

**repeat\_region**  
 9430..9506  
 /rpt\_family="ID4"

**repeat\_region**  
 9582..9652  
 /rpt\_family="(TG)n"

**repeat\_region**  
 9736..10073  
 /rpt\_family="ORR1A2"

**repeat\_region**  
 10088..10153  
 /rpt\_family="(CA)n"

**repeat\_region**  
 10526..10673  
 /rpt\_family="B1\_MM"

**Alignment Scores:**

Pred. No.: 5.04e+03 Length: 192682  
 Score: 45.00 Matches: 8  
 Percent Similarity: 88.89% Conservative: 0  
 Best Local Similarity: 88.89% Mismatches: 1  
 Query Match: 90.00% Indels: 0  
 DB: 10 Gaps: 0

US-10-014-101b-40 (1-9) x AC107769 (1-192682)

**OY** 1 ValLeuGlyGlyLeuGlyGlnPheCys 9

**DB** 20391 GTGTTAGGGGCTTGGGGGATTTTGT 20365

**RESULT 10**

**AC105345**

**LOCUS**

**DEFINITION**

**AC105345**

**VERSION**

**AC105345.3**

**KEYWORDS**

**HTG**

**SOURCE**

**ORGANISM**

**Homosapiens**

**REFERENCE**

**1 (bases 1 to 196902)**

**AUTHORS**

**Toward a complete human genome sequence**

**JOURNAL**

AC105345 196902 bp DNA linear PRI 03-JUL-2002  
 Homo sapiens BAC clone RP11-689P11 from 4, complete sequence.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 196902)

Sulston, J.E. and Waterston, R.

Toward a complete human genome sequence

Genome Res. 8 (11), 1097-1108 (1998)

MEDLINE	99063792	repeat_region	12321..12403
PUBMED	9847074	repeat_region	/rpt family="MER1_type"
REFERENCE	2 (bases 1 to 196902)	repeat_region	12402..12543
AUTHORS	Du,H., Haglund,K. and Spalding,L.	repeat_region	/rpt family="MER1_type"
TITLE	The sequence of Homo sapiens BAC clone RP11-689P11	repeat_region	13054..13410
JOURNAL	Unpublished (2001)	repeat_region	/rpt family="L1"
REFERENCE	3 (bases 1 to 196902)	repeat_region	14877..14987
AUTHORS	Waterston,R.H.	repeat_region	/rpt family="MIR"
TITLE	Direct Submission	repeat_region	15657..15956
JOURNAL	Submitted (31-DEC-2001) Genome Sequencing Center, Washington	repeat_region	/rpt family="Alu"
	University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA	repeat_region	19981..20035
REFERENCE	4 (bases 1 to 196902)	repeat_region	/rpt family="L2"
AUTHORS	Waterston,R.H.	repeat_region	20208..20290
TITLE	Direct Submission	repeat_region	/rpt family="L1"
JOURNAL	Submitted (06-MAR-2002) Genome Sequencing Center, Washington	repeat_region	20364..20502
	University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA	repeat_region	/rpt family="L2"
REFERENCE	5 (bases 1 to 196902)	repeat_region	23610..23744
AUTHORS	Waterston,R.	repeat_region	/rpt family="L2"
TITLE	Direct Submission	repeat_region	24921..25224
JOURNAL	Submitted (03-JUL-2002) Department of Genetics, Washington	repeat_region	/rpt family="Alu"
	University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA	repeat_region	27502..28052
COMMENT	On Jul 3, 2002 this sequence version replaced gi:19073865.	repeat_region	/rpt family="MER2_type"
FEATURES	Location/Qualifiers	repeat_region	28053..28267
	1..196902	repeat_region	/rpt family="Alu"
	/organism="Homo sapiens"	repeat_region	28273..28895
	/mol type="genomic DNA"	repeat_region	/rpt family="ERV1"
	/db xref="taxon:9606"	repeat_region	28896..29682
	/chromosomes="4"	repeat_region	/rpt family="ERV1"
	/maps="4"	repeat_region	29683..29764
	/clone="RP11-689P11"	repeat_region	/rpt family="ERV1"
	/clone lib="RPC1-11"	repeat_region	29771..29808
	636..941	repeat_region	/rpt family="Alu"
repeat_region	/rpt family="Alu"	repeat_region	/rpt family="MER2_type"
repeat_region	1539..1669	repeat_region	30583..31026
repeat_region	3351..3411	repeat_region	/rpt family="ERV1"
repeat_region	3468..3532	repeat_region	31027..31402
repeat_region	3758..3946	repeat_region	/rpt family="ERV1"
repeat_region	/rpt family="L1"	repeat_region	31401..31604
repeat_region	3970..5081	repeat_region	/rpt family="ERV1"
repeat_region	5089..5240	repeat_region	31621..32281
repeat_region	/rpt family="Alu"	repeat_region	/rpt family="ERV1"
repeat_region	5246..6635	repeat_region	36446..36893
repeat_region	6636..6938	repeat_region	/rpt family="ERV1"
repeat_region	6939..7100	repeat_region	/rpt family="MER2_type"
repeat_region	7101..7391	repeat_region	37495..37633
repeat_region	7392..7465	repeat_region	/rpt family="ERV1"
repeat_region	7472..7640	repeat_region	37634..37961
repeat_region	/rpt family="MIR"	repeat_region	/rpt family="Alu"
repeat_region	7916..8081	repeat_region	37962..38434
repeat_region	8527..8686	repeat_region	/rpt family="ERV1"
repeat_region	8749..9034	repeat_region	40158..40270
repeat_region	9039..9331	repeat_region	/rpt family="ERV1"
repeat_region	10313..10607	repeat_region	40271..40557
repeat_region	10609..10921	repeat_region	/rpt family="MER2_type"
repeat_region	/rpt family="Alu"	repeat_region	40558..40709
repeat_region	11805..11866	repeat_region	/rpt family="ERV1"
repeat_region	/rpt family="L1"	repeat_region	40710..41236
		repeat_region	/rpt family="ERV1"
		repeat_region	41237..41801
		repeat_region	/rpt family="ERV1"
		repeat_region	41802..41906
		repeat_region	/rpt family="MER2_type"
		repeat_region	41911..42256
		repeat_region	/rpt family="MER2_type"
		repeat_region	42279..42859
		repeat_region	/rpt family="ERV1"
		repeat_region	42968..43092
		repeat_region	/rpt family="ERV1"
		repeat_region	43093..43154
		repeat region	/rpt family="MER2_type"
			43432..43591

```

repeat_region      /rpt_family="MIR"
43626..43824
repeat_region      /rpt_family="MERI_type"
43905..44530
repeat_region      /rpt_family="ERV1"
44962..44988
repeat_region      /rpt_family="(T)n"
45005..45045
repeat_region      /rpt_family="AT_rich"
45124..45547
repeat_region      /rpt_family="ERV1"
45548..45854
repeat_region      /rpt_family="Alu"
45855..45988
repeat_region      /rpt_family="ERV1"
46243..46503
repeat_region      /rpt_family="L1"
46511..46736
repeat_region      /rpt_family="MaLR"
47703..47759
repeat_region      /rpt_family="(GGGTG)n"
47843..48117
repeat_region      /rpt_family="L1"
48135..48392
repeat_region      /rpt_family="Alu"
48394..48708
repeat_region      /rpt_family="Alu"
48839..49128
repeat_region      /rpt_family="Alu"
49319..49366
repeat_region      /rpt_family="L2"
49545..49775
repeat_region      /rpt_family="MIR"
49846..49906
repeat_region      /rpt_family="polypyrimidine"
49954..50079
repeat_region      /rpt_family="MIR"
50352..51081
repeat_region      /rpt_family="MaLR"
51261..51325
repeat_region      /rpt_family="MIR"

```

```

Alignment Scores:
Pred. No.:      5.14e+03      Length:      196902
Score:          45.00         Matches:      8
Percent Similarity: 88.89%   Conservative: 0
Best Local Similarity: 88.89% Mismatches:      1
Query Match:     90.00%      Indels:       0
DB:              9           Gaps:         0

```

US-10-014-101B-40 (1-9) x AC105345 (1-196902)

```

Qy      1 ValLeuGlyLeuGlyGlnPheCys 9
Db      97321 GTGTTGGGGGGCTTGGTAGCTTCTGC 97347

```

```

RESULT 11
AC007104      200000 bp      DNA      linear      HTG 17-JUL-1999
LOCUS      Homo sapiens chromosome 4, *** SEQUENCING IN PROGRESS ***, 16
DEFINITION      unordered pieces.
ACCESSION      AC007104
VERSION      AC007104.4 GI:5523795
KEYWORDS      HTG; HTGS_PHASE1.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE      1 (bases 1 to 200000)
AUTHORS      Stone,N.E., Schmutz,J.J., Cox,D.R. and Myers,R.M.
TITLE      Direct Submission
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 200000)

```

AUTHORS  
TITLE  
JOURNAL

COMMENT

Stone,N.E., Schmutz,J.J., Cox,D.R. and Myers,R.M.  
Direct Submission  
Submitted (17-MAR-1999) Department of Genetics, Stanford Human  
Genome Center, 855 California Avenue, Palo Alto, CA 94304, USA  
On Jul 17, 1999 this sequence version replaced gi:4757683.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 16 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence.  
\* as soon as it is available and the accession number will  
\* be preserved.

```

*      1      1369: contig of 1369 bp in length
*      1370      3774: gap of unknown length
*      3775      4878: contig of 1104 bp in length
*      4879      7283: gap of unknown length
*      7284      8555: contig of 1272 bp in length
*      8556      10960: gap of unknown length
*      10961      12207: contig of 1247 bp in length
*      12208      14612: gap of unknown length
*      14613      16262: contig of 1650 bp in length
*      16263      18667: gap of unknown length
*      18668      20176: contig of 1509 bp in length
*      20177      22581: gap of unknown length
*      22582      24279: contig of 1697 bp in length
*      24279      26684: gap of unknown length
*      26684      28199: contig of 1515 bp in length
*      28199      30604: gap of unknown length
*      30604      32360: contig of 1756 bp in length
*      32360      34765: gap of unknown length
*      34765      39959: contig of 5194 bp in length
*      39959      42364: gap of unknown length
*      42364      57085: contig of 14722 bp in length
*      57086      59490: gap of unknown length
*      59491      74906: contig of 15415 bp in length
*      74906      84274: gap of unknown length
*      84274      86679: contig of 6963 bp in length
*      86679      118061: gap of unknown length
*      118062      120466: contig of 31383 bp in length
*      120467      152965: gap of unknown length
*      152966      155371: contig of 32499 bp in length
*      155371      200000: gap of unknown length
*      200000      44630: contig of 44630 bp in length.

```

FEATURES  
source  
1..200000  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="4"

ORIGIN

```

Alignment Scores:
Pred. No.:      5.22e+03      Length:      200000
Score:          45.00         Matches:      8
Percent Similarity: 88.89%   Conservative: 0
Best Local Similarity: 88.89% Mismatches:      1
Query Match:     90.00%      Indels:       0
DB:              2           Gaps:         0

```

US-10-014-101B-40 (1-9) x AC007104 (1-200000)

```

Qy      1 ValLeuGlyLeuGlyGlnPheCys 9
Db      137654 GTGTTGGGGGGCTTGGTAGCTTCTGC 137680

```

RESULT 12  
AC116643  
LOCUS  
DEFINITION Homo sapiens BAC clone RP11-637J21 from 4, complete sequence.  
ACCESSION AC116643  
VERSION AC116643.11 GI:28173293  
KEYWORDS HTG.

linear  
DNA  
PRI 07-FEB-2003

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Sulston, J.E. and Waterston, R.

TITLE Toward a complete human genome sequence

JOURNAL Genome Res. 8 (11), 1097-1108 (1998)

MEDLINE 99063792

PUBMED 9847074

REFERENCE 2 (bases 1 to 214284)

AUTHORS Du, F. and Haglund, K.

TITLE The sequence of Homo sapiens BAC clone RP11-637J21

JOURNAL Unpublished (2001)

REFERENCE 3 (bases 1 to 214284)

AUTHORS Waterston, R.H.

TITLE Direct Submission

JOURNAL Submitted (30-MAR-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 4 (bases 1 to 214284)

AUTHORS Waterston, R.H.

TITLE Direct Submission

JOURNAL Submitted (29-JAN-2003) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 5 (bases 1 to 214284)

AUTHORS Waterston, R.H.

TITLE Direct Submission

JOURNAL Submitted (07-FEB-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

On Jan 31, 2003 this sequence version replaced gi:27753869.

----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc>

Contact: [sapiens@wustl.edu](mailto:sapiens@wustl.edu)

----- Summary Statistics

----- Center project name: H\_NH0637J21

-----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:

all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RP11-11 human BAC library was made from the blood of one male donor, as described by Ooegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tateno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong

and coworkers at <http://www.chori.org>  
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

Discrepant bases between AC007104 and clone sequence.

Data from AC105345 and AC007104 was used to finish this clone.

#### FEATURES

##### source

1. .214284  
/location="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="4"  
/map="4"

/clone="RP11-637J21"

/clone\_lib="RP11-11"

207. .514

/rpt\_family="Alu"

515. .622

/rpt\_family="L2"

709. .927

/rpt\_family="L2"

952. .1046

/rpt\_family="L2"

1311. .1741

/rpt\_family="ERV1"

1746. .2055

/rpt\_family="Alu"

2734. .2862

/rpt\_family="MIR"

3899. .4199

/rpt\_family="Alu"

5071. .5180

/rpt\_family="MIR"

5330. .5678

/rpt\_family="MALR"

5867. .6558

/rpt\_family="L1"

6635. .6946

/rpt\_family="Alu"

6963. .7046

/rpt\_family="L1"

7114. .7194

/rpt\_family="MIR"

7188. .7297

/rpt\_family="L2"

7851. .7927

/rpt\_family="MIR"

7928. .7958

/rpt\_family="(TCAA)n"

8945. .9131

/rpt\_family="MIR"

9216. .9330

/rpt\_family="L2"

9367. .9532

/rpt\_family="L1"

9533. .9850

/rpt\_family="Alu"

10243. .10351

/rpt\_family="MER1\_type"

11570. .11679

/rpt\_family="MIR"

12216. .12453

/rpt\_family="L1"

12454. .12597

/rpt\_family="Alu"

12792. .12947

/rpt\_family="L1"

12958. .13341

/rpt\_family="L1"

13355. .13659

/rpt\_family="L1"

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

```

repeat_region /rpt family="Alu"
13660..13793
repeat_region /rpt family="Alu"
13794..14175
repeat_region /rpt family="L1"
14176..14531
repeat_region /rpt family="MaLR"
14532..15069
repeat_region /rpt family="L1"
15070..15202
repeat_region /rpt family="Alu"
15203..15491
repeat_region /rpt family="Alu"
15492..15514
repeat_region /rpt family="Alu"
15515..16246
repeat_region /rpt family="L1"
16249..16578
repeat_region /rpt family="Alu"
16598..16657
repeat_region /rpt family="L1"
16658..16984
repeat_region /rpt family="Alu"
16985..17354
repeat_region /rpt family="L1"
17365..17435
repeat_region /rpt family="Alu"
17436..17782
repeat_region /rpt family="ERV1"
17784..17974
repeat_region /rpt family="MaLR"
18131..18439
repeat_region /rpt family="Alu"
18779..19085
repeat_region /rpt family="Alu"
19432..19542
repeat_region /rpt family="Tip100"
19543..20119
repeat_region /rpt family="ERV1"
20120..20505
repeat_region /rpt family="MaLR"
20506..20748
repeat_region /rpt family="ERV1"
20779..20884
repeat_region /rpt family="Tip100"
20889..21097
repeat_region /rpt family="Tip100"
21359..21466
repeat_region /rpt family="L2"

```

```

Alignment Scores:
Pred. No.: 5.56e+03 Length: 214284
Score: 45.00 Matches: 8
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 1
Query Match: 90.00% Indels: 0
DB: 9 Gaps: 0

```

US-10-014-101B-40 (1-9) x AC116643 (1-214284)

```

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9
Db 34363 GTCTGGGGGGGCTTGGTAGCTTCGC 34389

```

```

RESULT 13
AL513345/c 217225 bp DNA linear ROD 17-NOV-2001
LOCUS Mouse DNA sequence from clone RP23-129N7 on chromosome 15, complete
DEFINITION sequence.
ACCESSION AL513345
VERSION AL513345.20 GI:17017730
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)

```

# ORGANISM REFERENCE AUTHORS TITLE JOURNAL

Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 217225)  
Dunn, M.  
Direct Submission  
Submitted (17-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
On Nov 20, 2001 this sequence version replaced gi:15722124.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest. The following  
abbreviations are used to associate primary accession numbers given  
in the feature table with their source databases: Em: EMBL; Swi,  
SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP  
database can be found at  
http://www.sanger.ac.uk/projects/C\_elegans/wormpep RP23-129N7 is  
from the RPCI-23 Mouse PAC Library  
constructed by the group of Pieter de Jong.  
For further details see http://www.chori.org/bacpac/home.htm  
VECTOR: pBACE3.6

## COMMENT

This sequence is the entire insert of clone RP23-129N7.  
Location/Qualifiers  
1..217225  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="15"  
/clone="RP23-129N7"  
/clone\_lib="RPCI-23"  
4712..4781  
/note="Sequence from uni-directional dGTP big dye  
terminator reads only."  
206582..206603  
/note="Sequence from uni-directional primer reads only."

## FEATURES

source  
1..217225  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="15"  
/clone="RP23-129N7"  
/clone\_lib="RPCI-23"  
4712..4781  
/note="Sequence from uni-directional dGTP big dye  
terminator reads only."  
206582..206603  
/note="Sequence from uni-directional primer reads only."

## ORIGIN

```

Alignment Scores:
Pred. No.: 5.63e+03 Length: 217225
Score: 45.00 Matches: 8
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 1
Query Match: 90.00% Indels: 0
DB: 10 Gaps: 0

```

US-10-014-101B-40 (1-9) x AL513345 (1-217225)

```

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9
Db 122508 GTTCTGGAGGTTTGGGGTCCTTTTGT 122482

```

## RESULT 14

```

AC119853
LOCUS Mus musculus chromosome 16, clone RP23-231E23, complete sequence.
DEFINITION AC119853
ACCESSION AC119853.9 GI:41015153
VERSION AC119853.9
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

## REFERENCE

1 (bases 1 to 236274)

Birren,B., Nusbaum,C. and Lander,E.

Mus musculus chromosome 16, clone RP23-231E23

Unpublished

2 (bases 1 to 236274)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Lakocue,K., Lamazares,R., Landers,T., Lehocsky,J., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,D., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission  
Submitted (02-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 236274)

Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission  
Submitted (11-DEC-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 236274)

Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,

TITLE  
JOURNAL

## COMMENT

Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission  
Submitted (17-JAN-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jan 17, 2004 this sequence version replaced gi:39725840.

All repeats were identified using RepeatMasker:  
Smit,A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information  
Center project name: L22516  
Center clone name: 231\_E\_23

## FEATURES

## source

## Location/Qualifiers

1..236274  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="16"  
/map="16"  
/clone="RP23-231E23"  
/clone\_lib="RPC1-23 Female Mouse BAC"

## repeat\_region

1..3263

/rpt\_family="MERVL"

## misc\_feature

1..45

/note="wgs\_end\_extension"

## misc\_feature

46..51

/note="clone boundary"

clone\_end:SP6

site:EcORI

## repeat\_region

3264..3367

/rpt\_family="Ricksha"

## repeat\_region

3369..3861

/rpt\_family="MERVL\_LTR"

## repeat\_region

3862..4031

/rpt\_family="B3"

complement(4033..4140)

## repeat\_region

4169..4578

/rpt\_family="LIM2"

## repeat\_region

4579..5607

/rpt\_family="MERVL"

## repeat\_region

5616..5723

/rpt\_family="MT2B"

## repeat\_region

5771..5887

/rpt\_family="LIM2"

## repeat\_region

5888..5953

/rpt\_family="(GAA)n"

## repeat\_region

5954..5979

/rpt\_family="LIM2"

## repeat\_region

5980..6113

/rpt\_family="B1\_MM"

## repeat\_region

6114..6616

/rpt\_family="LIM2"

## repeat\_region

complement(7694..8299)

## repeat\_region

complement(8703..8887)

## repeat\_region

complement(11228..11501)

## repeat\_region

11564..12644

/rpt\_family="RLTR13D"

## repeat\_region

complement(12894..13095)

## repeat\_region

13100..13123

/rpt\_family="B3"

## repeat\_region

13136..13530

/rpt\_family="(T)n"

```

repeat_region      /rpt_family="MTA"
13531. .14628
/rpt_family="MT-int"
14629. .15023
/rpt_family="MTA"
15783. .16103
/rpt_family="MURVY-int"
16381. .16655
/rpt_family="MURVY-int"
17116. .18775
/rpt_family="MURVY-int"
19088. .19564
/rpt_family="IAPLTR2_MM"
19614. .19825
/rpt_family="MURVY-int"
complement(20880..21254)
/rpt_family="RMER17C"
22067. .22291
/rpt_family="B4A"
complement(22606..23019)
/rpt_family="MTC"
23597. .23666
/rpt_family="RMER6A"
23667. .23865
/rpt_family="RMER6A"
23892. .23936
/rpt_family="(TC)n"
23937. .24045
/rpt_family="RMER19B"
complement(24419..25145)
/rpt_family="ORR1A-int"
complement(25940..26264)
/rpt_family="MT-int"
complement(26570..26958)
/rpt_family="L1_MM"
complement(26993..27042)
/rpt_family="B1_MM"
27055. .27192
/rpt_family="RLTR14"
27345. .27493
/rpt_family="B1_MM"
complement(27799..27921)
/rpt_family="ID_B1"
complement(28034..28158)
/rpt_family="MLTC"

Alignment Scores:
Pred. No.:      6.09e+03      Length:      236274
Score:          45.00        Matches:      8
Percent Similarity: 88.89%    Conservative: 0
Best Local Similarity: 88.89% Mismatches:      1
Query Match:     90.00%      Indels:        0
DB:              10         Gaps:          0

US-10-014-101B-40 (1-9) x AC119853 (1-236274)
QY  1 ValLeuGlycylLeuGlyGlnPheCys 9
|||||
Db  51371 GTCTAGGGGTTGGGGGATTTTCT 51397
|||||

RESULT 15
AC124136/c
LOCUS      AC124136
DEFINITION Mus musculus clone RP23-250H5, WORKING DRAFT SEQUENCE, 8 unordered
pieces.
ACCESSION AC124136
VERSION   AC124136.3 GI:30017995
KEYWORDS  HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
           Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 277408)

```

# AUTHORS TITLE JOURNAL REFERENCE AUTHORS

Birren, B., Nusbaum, C. and Lander, E.  
Mus musculus, clone RP23-250H5  
Unpublished  
2 (bases 1 to 277408)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J. S., Dodge, S., Farol, S., Ferreira, P., Fitzgerald, M., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (09-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 277408)

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArelano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Farol, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (23-APR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Apr 17, 2003 this sequence version replaced gi:28195473.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L23602  
Center clone name: 250\_H\_5  
-----

\* NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will

# TITLE JOURNAL COMMENT

Submitted (23-APR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Apr 17, 2003 this sequence version replaced gi:28195473.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

\* be preserved.

1 67323: contig of 67323 bp in length

67324: gap of 100 bp

67424: contig of 1202 bp in length

68625: gap of 100 bp

68726: contig of 8745 bp in length

77470: gap of 100 bp

77570: contig of 77311 bp in length

154881: gap of 100 bp

154882: contig of 14688 bp in length

169670: gap of 100 bp

169769: contig of 28120 bp in length

197770: gap of 100 bp

197890: contig of 18228 bp in length

269818: gap of 100 bp

269918: contig of 7491 bp in length.

# FEATURES

source 1..277408  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone="RP23-250H5"  
/clone\_lib="RPCI-23 Female Mouse BAC"  
1..67323  
/notes="assembly\_fragment  
clone\_end:SP6  
vector\_side:left"  
67424..68625  
/notes="assembly\_fragment"  
68726..77470  
/notes="assembly\_fragment"  
77571..154881  
/notes="assembly\_fragment"  
154982..169669  
/notes="assembly\_fragment"  
169770..197889  
/notes="assembly\_fragment"  
197990..269817  
/notes="assembly\_fragment"  
269918..277408  
/notes="assembly\_fragment  
clone\_end:T7  
vector\_side:right"

## ORIGIN

Alignment Scores:  
Pred. No.: 7.06e+03 Length: 277408  
Score: 45.00 Matches: 8  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 1  
Query Match: 90.00% Indels: 0  
DB: 2 Gaps: 0

US-10-014-101B-40 (1-9) x AC124136 (1-277408)

Qy 1 ValLeuGlyLeuGlyGlnPheCys 9  
|||||  
Db 218386 GTGTAGCGGTTTGGGGGATTTTGT. 218360

## RESULT 16

E07190/c E07190 1509 bp DNA linear PAT 29-SEP-1997  
DEFINITION DNA encoding carotenoid synthetase from *Thermus thermophilus*.  
ACCESSION E07190  
VERSION E07190.1 GI:2175331  
KEYWORDS JP 1994098774-A/1.  
SOURCE *Thermus aquaticus*  
ORGANISM *Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae; Thermus.*  
REFERENCE 1 (bases 1 to 1509)  
AUTHORS Hoshino, T. and Nakahara, T.  
TITLE GENE DNA FRAGMENT ORIGINATED FROM *THERMUS* GENUS BACTERIA AND

## JOURNAL

## COMMENT

PARTICIPATING IN BIOSYNTHESIS OF CAROTENOID AND ITS USE  
Patent: JP 1994098774-A 1 12-APR-1994;  
MITSUBISHI PETROCHEM CO LTD  
OS *Thermus thermophilus*  
PN JP 1994098774-A/1  
PD 12-APR-1994  
PF 25-SEP-1992 JP 1992256580  
PI HOSHINO TAKAYUKI, NAKAHARA TADAATSU  
PC C12N15/31.C12N1/20.C12N1/21.C12P23/00.(C12N15/31.C12R1:01), PC  
(C12N1/20).

PC C12R1:01.(C12N1/21.C12R1:01).(C12P23/00.C12R1:01); CC  
strandedness: Double;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: No;  
FH Key Location/Qualifiers  
FT source 1..1509  
FT /organism="Thermus thermophilus" FT  
FT /strain="HB27" /clones="pCOP1.pCOP2"  
FT 5'UTR 1..59  
FT CDS 60..1103  
FT /product="carotenoid biosynthesis from FT  
FT *Thermus thermophilus*  
FT 3'UTR 1104..1509.  
FT Location/Qualifiers  
FT 1..1509  
FT /organism="Thermus aquaticus"  
FT /mol\_type="genomic DNA"  
FT /db\_xref="taxon:271"

## FEATURES

source 1..1509  
Location/Qualifiers  
1..1509  
/organism="Thermus aquaticus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:271"

## ORIGIN

Alignment Scores:  
Pred. No.: 90.1 Length: 1509  
Score: 44.00 Matches: 7  
Percent Similarity: 88.89% Conservative: 1  
Best Local Similarity: 77.78% Mismatches: 1  
Query Match: 88.00% Indels: 0  
DB: 6 Gaps: 0

US-10-014-101B-40 (1-9) x E07190 (1-1509)

Qy 1 ValLeuGlyLeuGlyGlnPheCys 9  
:::|||||  
Db 61 ATTCTAGCGGCTCGGGGACTTTTGC 35

## RESULT 17

## LOCUS

## DEFINITION

ALL139349 68743 bp DNA linear PRI 22-FEB-2001  
Human DNA sequence from clone RP11-261P9 on chromosome 20 Contains  
ESTs, STSs, GSSs and a CpG island. Contains the 3' part of the  
STX16 gene for syntaxin 16 and a novel gene for a protein similar  
to fly CG7340 and human putative aminopeptidase ZK353.6 in  
chromosome 3 (EC 3.4.11.-), complete sequence.

ACCESSION ALL139349.36 GI:11544447

VERSION ALL139349

KEYWORDS HTG; aminopeptidase; CG7340; CpG island; syntaxin 16.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 68743)

AUTHORS Moore, M

TITLE Direct Submission

JOURNAL Submitted (20-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

REQUESTS: clonesrequest@sanger.ac.uk

On Dec 4, 2000 this sequence version replaced gi:11414473.

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at

[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human

chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/chr20>

IMPORTANT: this sequence is not the entire insert of clone RP11-261p9 it may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.

The true right end of clone RP11-261p9 is at 68743 in this sequence. The true left end of clone RP4-806W20 is at 51526 in this

sequence. The true right end of clone RP5-907D15 is at 100 in this sequence. RP11-261p9 is from the library RPCI-11.1 constructed by

the group of Pieter de Jong. For further details see

<http://www.chori.org/bacpac/home.htm>

VECTOR: pBACe3.6

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

## FEATURES

source

1..68743  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="20"  
/clone="RP11-261p9"  
/clone\_lib="RPCI-11.1"  
719..894

repeat\_region

repeat\_region

repeat\_region

gene

mRNA

mRNA

/evidence=not\_experimental  
Join(<2765..2872,3262..3402,5787..5878,6429..6572,  
8906..8986,11462..14775)  
/gene="STX16"

/product="BA261P9.1.3 (syntaxin 16C)"

/note=match: CDNAS: Em:AF008937 Em:AL359056 Em:AK026974

match: ESTs: Em:BF096123 Em:BE281441 Em:AV746848

Em:BE781705 Em:BE536701 Em:BE305443 Em:BE881751

Em:BE76318 Em:AI733726 Em:BE290417 Em:AL120975

Em:AA100145 Em:J01883 Em:AL118944 Em:AV683718 Em:AA789674

Em:AV725373 Em:AV709337 Em:AW382368 Em:BE147200

Em:AW545644 Em:BE689158 Em:AA763463 Em:BE019106

Em:BE937266 Em:AW606511 Em:AA844452 Em:BE684240

Em:AA676370 Em:BE738350 Em:RA78673 Em:AI632175 Em:BF057674

Em:AI651286 Em:AL047428 Em:R97389 Em:AW977869 Em:AL047429

Em:BE686410 Em:AI338386 Em:BE246533 Em:AA057475 Em:N40442

Em:AA137712 Em:AV750362 Em:AW473361 Em:AI278678

Em:AA203630 Em:BE501312 Em:BE321514 Em:BE0003924

Em:AA906113 Em:W28230 Em:T30587 Em:BB314862 Em:R94627

Em:BB413687 Em:R97872 Em:AA459475 Em:AW206023 Em:AI356379

Em:H25921 Em:BE866818 Em:AV315580 Em:BF204294 Em:BB541020

Em:AI567688 Em:AW290881 Em:AV343686 Em:BB155303

Em:BB310914 Em:AW376257 Em:R63084 Em:BF204735 Em:T97621

Em:AI016616 Em:W40487 Em:AW958296 Em:W40488 Em:BE161763

Em:AI457303 Em:W40537 Em:AA872491 Em:BB153094 Em:W40536

Em:AA249864 Em:AI873223 Em:AV728943 Em:BE753527

Em:AI956006 Em:AA828279 Em:AI989398 Em:AI288570

Em:BE782754 Em:AI871762 Em:Z36994 Em:BE843000 Em:AW027820

Em:AI276677 Em:AI868672 Em:H43754 Em:AV747611 Em:AW172483

Em:AV727930 Em:AW785218 Em:AA776032 Em:AA459250

Em:AV659728 Em:AV315794 Em:AA594117 Em:AA194704

Em:AI334818 Em:AI027633 Em:AA458607 Em:AA629872 Em:D20096

Em:AI249473 Em:BE168567 Em:AV329041 Em:BE737892

Em:AI890882 Em:AI74173 Em:N30151 Em:AW612089 Em:AW475172

Em:AV747307 Em:AI493393 Em:AA203345 Em:L44457 Em:H63081

Em:AA319231 Em:AW374052 Em:N35763 Em:BB462304 Em:BB162451

Em:AA076062 Em:AI370022 Em:BE674002 Em:AW972891

Em:AW964476 Em:AA725590 Em:AA164684 Em:AA165099

Em:AA236379 Em:AI655600 Em:BB416468 Em:BE178290

Em:AI333339 Em:BE836408 Em:BE836409 Em:H63165 Em:BF058685

Em:AI739076 Em:AV362504 Em:BE735967 Em:AW542371

Em:BE927688 Em:AA311035 Em:N44527 Em:AA682368 Em:BE220136

Em:R80312 Em:AW382369 Em:BB315712 Em:AA236380 Em:AA057371

Em:AA814622 Em:AW874573 Em:AA365802 Em:W42617 Em:W42618  
 Em:W84550 Em:AA831078 Em:BE836411 Em:BE205862 Em:AW653492  
 Em:AA167527 Em:BE836412 Em:AA167528 Em:AI064903  
 Em:BE836415 Em:BE548226 Em:N24793 Em:T97573 Em:BE856414  
 Em:AI334393 Em:BE245863 Em:AW130335 Em:AV114038  
 Em:AI083695 Em:AI829831 Em:AW378653 Em:N24133 Em:AV721905  
 Em:AW553605 Em:AW505186 Em:N33763 Em:R63128 Em:W42621  
 Em:BE261034 Em:W42622 Em:AA812137 Em:AA524825 Em:AA354822  
 Em:BE378212 Em:AW891599 Em:Z41983 Em:AI911671 Em:AA091987  
 Em:AA489328 Em:F20270 Em:AA319192 Em:BE927706 Em:AI356428  
 Em:AV302442 Em:BE878126 Em:AI823467 Em:AV702161  
 Em:AA256336 Em:BE937267 Em:AI570829 Em:AI873185  
 Em:BF156017 Em:AA036755 Em:AV724421 Em:AW504146  
 Em:AA256337 Em:AI949076 Em:AW542361 Em:AI366948  
 /evidence=not\_experimental  
 join(-2765..2872,3262..3402,4566..4728,5787..5878,  
 6429..6572,8906..8986,11462..11566)  
 /gene="STX16"  
 /note="isoforms 1 and 2  
 continues in dJ907D15 (AL050327)  
 match: proteins: Tr:O14661 Tr:Q9SBD4 Tr:O04258 Tr:Q9VR90  
 Tr:O04639 Tr:Q9SUJ1 Tr:O65359 Tr:Q9P6P1 Tr:Q9S564  
 Tr:O70319 Tr:O88385 Tr:Q9TY08 Sw:O15400 Tr:Q9VU45"  
 /codon\_start=1  
 /evidence=not\_experimental  
 /product="bA261P9.1.1 (syntaxin 16A / syntaxin 16B)"  
 /protein\_id="CAC19812.1"  
 /db\_xref="GI:12044552"  
 /db\_xref="GOA:O14662"  
 /db\_xref="UniProt/Swiss-Prot:O14662"  
 /translation="LADDMALVSGISLDPAAIGVTKRPKPPKWDGVDEIQYDVGR1

## CDS

## Alignment Scores:

Pred. No.: 3.05e+03 Length: 68743  
 Score: 44.00 Matches: 8  
 Percent Similarity: 88.89% Conservative: 0  
 Best Local Similarity: 88.89% Mismatches: 1  
 Query Match: 88.00% Indels: 0  
 DB: 9 Gaps: 0

US-10-014-101b-40 (1-9) x AL139349 (1-68743)

QY 1 valLeuGlyClyLeuGlyGlnPheCys 9

DB 14835 GTTAGGAGGAGGCTGGCCAGTTTGT 14861

## RESULT 18

AC006198 AC006198 163410 bp DNA linear PRI 19-DEC-1998  
 LOCUS Homo sapiens, clone hRPK.3\_A\_1, complete sequence.  
 DEFINITION  
 AC006198  
 VERSION  
 AC006198.1 GI:4038072  
 KEYWORDS  
 HTG.  
 SOURCE  
 Homo sapiens (human)

ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
 1 (bases 1 to 163410)

AUTHORS  
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.

## TITLE

## JOURNAL

## REFERENCE

2 (bases 1 to 163410)  
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,  
 Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,  
 Castle,A., Cerny,J., Collangelo,M., Collins,S., Collymore,A.,  
 Cooke,P., Depayre,E., Devon,K., Dewar,K., Donelan,L., Ferreira,P.,  
 Fitzhugh,W., Forrest,C., Funke,R., Gage,D., Galagan,J., Gardyna,S.,  
 Grant,G., Hagos,B., Heaford,A., Herena,L., Horton,L., Howland,J.C.,  
 Jones,C., Kann,L., Karatas,A., Lehoczy,J., Macdonald,P.,  
 Marquis,N., McSwan,P., McGurk,A., McKernan,K., Meldrim,J.,  
 Molla,M., Morris,W., Morrow,J., Mychaleckyj,J., Naylor,J.,  
 Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B., Peterson,K.,  
 Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,

Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,  
 Tesfaye,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,  
 Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.  
 Direct Submission  
 Submitted (09-DEC-1998) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 163410)  
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,  
 Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,  
 Castle,A., Cerny,J., Collangelo,M., Collins,S., Collymore,A.,  
 Cooke,P., DeArrellano,K., Depayre,E., Devon,K., Dewar,K.,  
 Donelan,L., Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R.,  
 Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A.,  
 Herena,L., Horton,L., Howland,J.C., Jones,C., Kann,L., Karatas,A.,  
 Lehoczy,J., Macdonald,P., Marquis,N., McSwan,P., McGurk,A.,  
 McKernan,K., Meldrim,J., Molla,M., Morris,W., Morrow,J.,  
 Mychaleckyj,J., Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P.,  
 Pavlin,B., Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A.,  
 Severy,P., Stange-Thomann,N., Stojanovic,N., Stone,C.,  
 Subramanian,A., Tesfaye,S., Torruella-Miller,I., Vassiliev,H.,  
 Vo,A., Wagner,A., Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.  
 Direct Submission  
 Submitted (19-DEC-1998) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Dec 19, 1998 this sequence version replaced gi:1983559.  
 All repeats were identified using RepeatMasker: Smt, A.F.A. &  
 Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html.

## FEATURES

## Source

1..163410  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="hRPK.3\_A\_1"  
 /clone\_lib="RPc1-11 human BAC library"  
 complement(863..922)  
 /rpt\_family="Tigger4(Zombi)"  
 complement(906..944)  
 /rpt\_family="Tigger4(Zombi)"  
 2836..3103  
 /rpt\_family="LTR16C"  
 3733..4019  
 /rpt\_family="AluSq"  
 4020..4045  
 /rpt\_family="(CAAA)n"  
 complement(4543..4906)  
 /rpt\_family="MLT1A2"  
 complement(6270..6588)  
 /rpt\_family="LIME3"  
 complement(6745..6939)  
 /rpt\_family="MER5A"  
 7414..7594  
 /rpt\_family="MER33"  
 complement(7595..7741)  
 /rpt\_family="L1Pall"  
 7742..7924  
 /rpt\_family="MER33"  
 8585..8768  
 /rpt\_family="MIR"  
 complement(8941..9389)  
 /rpt\_family="L2"  
 complement(10056..10565)  
 /rpt\_family="LIMASA"  
 complement(10640..10689)  
 /rpt\_family="5S"  
 complement(11300..11645)  
 /rpt\_family="THE1B"  
 complement(12213..12257)  
 /rpt\_family="(CA)n"  
 13319..13345  
 /rpt\_family="AT\_rich"  
 13353..13387  
 /rpt\_family="MER5A"

```

repeat_region complement(14301..14453)
/rpt family="MIR"
repeat_region complement(15130..15181)
/rpt family="POLY A"
repeat_region complement(15182..15498)
/rpt family="LTR17"
repeat_region complement(15489..15637)
/rpt family="HERV17"
repeat_region complement(15855..15903)
/rpt family="L2"
repeat_region 15304..15887
/rpt family="LTR5"
repeat_region complement(16888..16979)
/rpt family="L2"
repeat_region 17191..17235
/rpt family="purine-rich"
repeat_region 17715..17853
/rpt family="MIR"
repeat_region 17936..18187
/rpt family="L1ME"
repeat_region 18298..18337
/rpt family="AT rich"
repeat_region complement(18338..18599)
/rpt family="MIR"
repeat_region 18791..18847
/rpt family="AT rich"
repeat_region 19370..19618
/rpt family="L1ME1"
repeat_region 19682..19778
/rpt family="AT rich"
repeat_region complement(22838..23169)
/rpt family="Trigger1"
repeat_region 23170..23470
/rpt family="AluSq"
repeat_region complement(23471..25315)
/rpt family="Trigger1"
repeat_region 25796..25833
/rpt family="(TA)n"
repeat_region 26875..26906
/rpt family="(TA)n"
repeat_region complement(27024..27141)
/rpt family="(CATA)n"
repeat_region 27207..27397
/rpt family="(TA)n"
repeat_region 27595..27887
/rpt family="L2"
repeat_region complement(28463..28533)
/rpt family="(CACG)n"
repeat_region 28574..28850
/rpt family="AluSx"
repeat_region 28851..28898
/rpt family="A-rich"
repeat_region 31114..31352
/rpt family="MLT1G"
repeat_region complement(31904..32203)
/rpt family="AluJo"
repeat_region 34165..34238
/rpt family="MLT1I"
repeat_region 34410..34563
/rpt family="MLT1H"
repeat_region complement(34834..35036)
/rpt family="MIR"
repeat_region complement(35322..35607)
/rpt family="L1MC4"
repeat_region 35705..35741
/rpt family="polypurine"
repeat_region complement(36027..36127)
/rpt family="MIR"
repeat_region 36211..36302
/rpt family="L2"
repeat_region 36545..37213
/rpt family="L2"
repeat_region complement(39886..40054)

```

```

/rpt family="MER5A"
40056..40407
/rpt family="THE1A"
complement(40682..40950)
/rpt family="L2"
41006..41321
/rpt family="MER33"
41475..41842
/rpt family="LTR16C"
complement(41861..42170)
/rpt family="L2"
complement(42271..42676)
/rpt family="L2"
complement(43808..44109)
/rpt family="AluSg"
46484..46664
/rpt family="MER58C"
47030..47336
/rpt family="MLT1F"
complement(48248..48961)
/rpt family="MER21B"
complement(48990..49095)
/rpt family="MIR"
complement(49658..49718)
/rpt family="L2"
50388..50412
/rpt family="(TAA)n"
complement(50521..50638)
/rpt family="MIR"

```

## Alignment Scores:

Pred. No.:	6.79e+03	Length:	163410
Score:	44.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	87.50%	Mismatches:	0
Query Match:	88.00%	Indels:	0
DB:	9	Gaps:	0

US-10-014-101b-40 (1-9) x AC006198 (1-163410)

QY 2 LeuGlyGlyLeuGlyGlnPheCys 9

Db 123956 ATAGTGGCTGGGGCAATCTGT 123979

## RESULT 19

```

AC105162
LOCUS AC105162 177725 bp DNA linear HTG 27-AUG-2004
DEFINITION Mus musculus chromosome 1 clone RP24-92011 map 1, *** SEQUENCING IN
PROGRESS ***, 4 ordered pieces.
ACCESSION AC105162
VERSION AC105162.9 GI:51572058
KEYWORDS HTG; HTGS PHASE2; HTGS FULLTOP; HTGS_ACTIVEFIN.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

```

## REFERENCE

```

AUTHORS 1 (bases 1 to 177725)
TITLE Birren,B., Nusbaum,C. and Lander,E.
JOURNAL Mus musculus chromosome 1, clone RP24-92011
REFERENCE 2 (bases 1 to 177725)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Girde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Katat,A., Karatas,A., Kells,C., LaRocque,K.,
Lamazares,R., Landers,T., Lechoczky,J., Levine,R., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,

```





```
repeat_region 22628..22755
/rpt_family="B1P"

Alignment Scores:
Pred. No.: 8.05e+03 Length: 196529
Score: 44.00 Matches: 7
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 87.50% Mismatches: 0
Query Match: 88.00% Indels: 0
DB: 10 Gaps: 0

US-10-014-101B-40 (1-9) x AC102630 (1-196529)

QY 2 LeuGlyClyLeuGlyGlnPheCys 9
|||:::|||||
Db 194599 CTGGAGGGATGGCGCAGTTTGT 194622

RESULT 21
AE017222/c 232605 bp DNA circular BCT 10-MAY-2004
LOCUS Thermus thermophilus HB27 plasmid pTT27, complete sequence.
DEFINITION AE017222
ACCESSION AE017222
VERSION AE017222.1 GI:46197920
KEYWORDS Thermus thermophilus HB27
SOURCE Thermus thermophilus HB27
ORGANISM Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
Thermus.
REFERENCE 1 (bases 1 to 232605)
AUTHORS Henne, A., Brueggemann, H., Raasch, C., Wierzer, A., Hartsch, T.,
Liesegang, H., Johann, A., Lienard, T., Gohl, O., Martinez-Arias, R.,
Jacobi, C., Starkuviene, V., Schlenczek, S., Dencker, S., Huber, R.,
Klenk, H.-P., Overbeek, R., Kramer, W., Merkl, R., Gottschalk, G. and
Fritz, H.-J.
TITLE The genome sequence of the extreme thermophile Thermus thermophilus
JOURNAL Nat. Biotechnol. 22 (5), 547-553 (2004)
PUBMED 15064768
REFERENCE 2 (bases 1 to 232605)
AUTHORS Henne, A., Brueggemann, H., Raasch, C., Wierzer, A., Hartsch, T.,
Liesegang, H., Johann, A., Lienard, T., Gohl, O., Martinez-Arias, R.,
Jacobi, C., Starkuviene, V., Schlenczek, S., Dencker, S., Huber, R.,
Klenk, H.-P., Overbeek, R., Kramer, W., Merkl, R., Gottschalk, G. and
Fritz, H.-J.
TITLE Direct Submission
JOURNAL Submitted (03-FEB-2004) Goettingen Genomics Laboratory, Institute
for Microbiology and Genetics, Grisebachstr. 8, Goettingen D-37077,
Germany
FEATURES
source
1. 232605 Location/Qualifiers
/organism="Thermus thermophilus HB27"
/mol_type="genomic DNA"
/strain="HB27"
/db_xref="taxon:262724"
/plasmid="pTT27"
complement(301..1623)
/locus_tag="TTP0001"
complement(301..1623)
/locus_tag="TTP0001"
/codon_start=1
/transl_table=11
/product="cobyrinic acid A,C-diamide synthase"
/protein_id="AAS82331.1"
/db_xref="GI:46197921"
/translation="MRLLAAHSHSGAGTKTTSIALLLALRARGLRVQPFKVPDPIIDP
THLEKAAARPNYLDGFFDETGLLALFRHARGADFALIEGVMGFLDCKDPRGEVGS
TAQVALLKAPVAVVDKMGAGSIAPLAGPDPHPGVVGVFANRVGSRHARIL
KEALKAVGFLPGLWLPQDPALELPERHGLVLAGVAPLEALRRAFRVVDLGSVLRLA
ASAPPLERAPFLPPRRPARVAYADKAFSYYPEALELLEALGAEILVPPSPLEDE
ALPKHALILGGYPBELFARLSANVALREAIRFPPTVAECGGMYLSQGLWGER
FFPMVGLVPGCEARMAERPVILGYREVALRDSVPARKQOAFKGHEFYARLPASPPAW
RRLGGEVEGYTDGTVLASFVHLYLPARPEGVERLUGLASRGKAAGAS"
complement(1623..2186)
/locus_tag="TTP0002"

gene
complement(1623..2186)
/locus_tag="TTP0002"
/codon_start=1
/transl_table=11
/product="cob(i)alamin adenosyltransferase"
/protein_id="AAS82332.1"
/db_xref="GI:46197922"
/translation="MEEPQRQPYAKPQQRGRLLVYTGDKGKSTAAFGALRAHG
RLKVRIFQFIKHGTARFEHRAFLSLGIPIELGLDGFWRSDRLARSALAEQWGR
AKELSLSGTDLVDLDEATFYPLRYGNVSLSEEFLEVLRAHPHVHVVTGCGAPEALLE
LADTVTEMRKVHAFDQGVPAQKQIEH"
complement(2218..2754)
/locus_tag="TTP0003"
complement(2218..2754)
/locus_tag="TTP0003"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAS82333.1"
/db_xref="GI:46197923"
/translation="MSEDRLLERLEKLEGVETTVAVLPPLIQDLSRRIDALREEVKA
ELRREARLEERFHLERKIQEVETKLESRIQNLRLGRQIQETRLGTRIQVEEA
KLDROIQTEAQLRIHIVAVETKLEGRIOVENRLEAQIALCALCQEMKABIGSAFTRVM
LYLTAVGVVLLALLSVFER"
complement(2759..3478)
/locus_tag="TTP0004"
complement(2759..3478)
/locus_tag="TTP0004"
/EC_number="2.4.2.21"
/codon_start=1
/transl_table=11
/product="cobalamin [5'-phosphate] synthase"
/protein_id="AAS82334.1"
/db_xref="GI:46197924"
/translation="MKAPRLALITLVLPLAPKGVGEEDFKRSVAFPLAGYLLGLPL
ALLALPLPPGLSALGVALGLGLDGLDLADALGAPREERLILKDPHL
GAFAPGVGVYLLILLFOALALVQDPLFLLLFPGARFAFLPFLHRYPLPGPMAALVR
GGPWFALLPPLFLLYPLPALALLAAWVARLAWARLGGINGDALGAMTALGEVV
LLLAQALGPAPSSRAGPLP"
complement(3475..4485)
/locus_tag="TTP0005"
complement(3475..4485)
/locus_tag="TTP0005"
/EC_number="2.4.2.21"
/codon_start=1
/transl_table=11
/product="nicotinate-nucleotide-dimethylbenzimidazole
phosphoribosyltransferase"
/protein_id="AAS82335.1"
/db_xref="GI:46197925"
/translation="MVDPEVFAQARLMDQLTKPPRALGYLBEVALRLAALQGRVKPE
LGRGAVVVAADHGVVAGVSAYPQEVTVQVNLNLSGGAIQFALTADCAVYVLDV
GVVGBLPDHPGLKRRKRGPTANLARGPMTPEEARALLAGRAARRAIAGATLLIA
AGDMIGNTTAAALTAALGIPPRVAVGRTGVGEEGLRRKREAVALARLHPCMG
PLEVAAEVGGLLELVAIGVYLEGYENGFLVLDGPPVTAGALLAKWMAQLRDLHLPAG
HLSRPFGRHHRHQUEALGLRFLDLADLGEGTGAVLAMPGLRAAARILHFWATFQAGVS
RG"
4509..5042
/locus_tag="TTP0006"
4509..5042
/locus_tag="TTP0006"
/codon_start=1
/transl_table=11
/product="phosphoglycerate mutase"
/protein_id="AAS82336.1"
/db_xref="GI:46197926"
/translation="WELMVRHGETLWNRGRLLGWTDPLTAEGEAQARRLKALPSS
LPAFSSDLLRARQTLAFLAGSPRLYPELRHIFGALGALWETLDPRIKALLRFQGF
HPGVSLSAQFERVATFRLGLEKAPAVLFTHGGVVRVLRALGEDGLVPPGSAVYDM
PRRVLRALDGEATG"
5307..5969
/locus_tag="TTP0007"
5307..5969
```

Alignment Scores:			
Pred. No.:	9.4e+03	Length:	232605
Score:	44.00	Matches:	7
Percent Similarity:	88.89%	Conservative:	1
Best Local Similarity:	77.78%	Mismatches:	1
Query Match:	88.00%	Indels:	0
DB:	1	Gaps:	0
US-10-014-101B-40 (1-9) x AE017222 (1-232605)			
QY	1 ValLeuGlyGlyLeuGlyGlnPheCys 9		
Db	51636 ATTCTAGCGCGCCTCGGGAGCTTTTGC 51610		
RESULT 22			
AY661656/c	AY661656	279448 bp	linear
LOCUS	Sorghum bicolor clone BAC 88M4, complete sequence.		
DEFINITION	AY661656		
ACCESSION	AY661656.1	GI:49359135	
VERSION	HTG.		
KEYWORDS	Sorghum bicolor (sorghum)		
SOURCE	Sorghum bicolor		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Clade; Panicoideae; Andropogoneae; Sorghum.		
REFERENCE	1 (bases 1 to 279448)		
AUTHORS	Islam-Faridi,M.N., Kim,J.-S., Klein,P.E., Stelly,D. Klein,R.R. and Mullet,J.E.		
TITLE	Cytogenetic Analysis of Sorghum Chromosome 3 and A Chromosome 1 Reveals Expansion of Pericentromeric Region in Sorghum		
JOURNAL	Unpublished (2004)		
REFERENCE	2 (bases 1 to 279448)		
AUTHORS	Klein,R.R., Klein,P.E., Mullet,J.E., Minx,P. and M. Direct Submission		
TITLE	Submitted (21-JUN-2004) USDA-ARS, Southern Plains Research Center, 2765 FM Road, College Station, TX Location/Qualifiers		
FEATURES	1. .279448		
source	/organism="Sorghum bicolor"		
	/mol_type="genomic DNA"		
	/cultivar="IS320C"		
	/db_xref="taxon:4558"		
	/chromosome="8"		
	/map="104.2-111.2 cm"		
	/clone="BAC 88M4"		
ORIGIN			
Alignment Scores:			
Pred. No.:	1.11e+04	Length:	279448
Score:	44.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	87.50%	Mismatches:	0
Query Match:	88.00%	Indels:	0
DB:	8	Gaps:	0
US-10-014-101B-40 (1-9) x AY661656 (1-279448)			
QY	2 LeuGlyGlyLeuGlyGlnPheCys 9		
Db	51903 TTGGAGGCGATCGGCAGTTTTGT 51880		
RESULT 23			
AR042369/c	AR042369	5183 bp	linear
LOCUS	Sequence 3 from patent US 5811304.		
DEFINITION	AR042369		
ACCESSION	AR042369.1	GI:5962865	
VERSION	Unknown.		
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unknown.		

Unclassified.  
REFERENCE 1 (bases 1 to 5183)  
AUTHORS Huang, S.  
TITLE Nucleic acid molecules encoding retinoblastoma protein-interacting zinc finger proteins  
JOURNAL Patent: US 5811304-A 3 22-SEP-1998;  
FEATURES Location/Qualifiers  
source 1..5183  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Alignment Scores: 441 Length: 5183  
Pred. No.: 43.00 Matches: 8  
Score: 88.89% Conservative: 0  
Percent Similarity: 88.89% Mismatches: 1  
Best Local Similarity: 88.89% Indels: 0  
Query Match: 86.00% Gaps: 0  
DB: 6

US-10-014-101B-40 (1-9) x AR042369 (1-5183)  
QY 1 ValLeuGlyLeuGlyGlnPheCys 9  
|||||  
Db 1608 GTTCTGGTGGCTGGGCGCTGTTCTGC 1582

RESULT 24  
AR052277/c  
LOCUS AR052277 5183 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 3 from patent US 5831008.  
ACCESSION AR052277  
VERSION AR052277.1 GI:5975641  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 5183)  
AUTHORS Huang, S.  
TITLE Retinoblastoma protein-interacting zinc finger proteins  
JOURNAL Patent: US 5831008-A 3 03-NOV-1998;  
FEATURES Location/Qualifiers  
source 1..5183  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Alignment Scores: 441 Length: 5183  
Pred. No.: 43.00 Matches: 8  
Score: 88.89% Conservative: 0  
Percent Similarity: 88.89% Mismatches: 1  
Best Local Similarity: 88.89% Indels: 0  
Query Match: 86.00% Gaps: 0  
DB: 6

US-10-014-101B-40 (1-9) x AR052277 (1-5183)  
QY 1 ValLeuGlyLeuGlyGlnPheCys 9  
|||||  
Db 1608 GTTCTGGTGGCTGGGCGCTGTTCTGC 1582

RESULT 25  
AR241188/c  
LOCUS AR241188 5868 bp DNA linear PAT 20-DEC-2002  
DEFINITION Sequence 3 from patent US 6468985.  
ACCESSION AR241188  
VERSION AR241188.1 GI:27286418  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 5868)  
AUTHORS Huang, S.  
TITLE Retinoblastoma protein-interacting zinc finger proteins

JOURNAL Patent: US 6468985-A 3 22-OCT-2002;  
FEATURES Location/Qualifiers  
source 1..5868  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Alignment Scores: 494 Length: 5868  
Pred. No.: 43.00 Matches: 8  
Score: 88.89% Conservative: 0  
Percent Similarity: 88.89% Mismatches: 1  
Best Local Similarity: 88.89% Indels: 0  
Query Match: 86.00% Gaps: 0  
DB: 6

US-10-014-101B-40 (1-9) x AR241188 (1-5868)  
QY 1 ValLeuGlyLeuGlyGlnPheCys 9  
|||||  
Db 1728 GTTCTGGTGGCTGGGCGCTGTTCTGC 1702

RESULT 26  
AR262645/c  
LOCUS AR262645 5868 bp DNA linear PAT 29-JAN-2003  
DEFINITION Sequence 3 from patent US 6323335.  
ACCESSION AR262645  
VERSION AR262645.1 GI:28074187  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 5868)  
AUTHORS Huang, S.  
TITLE Retinoblastoma protein-interacting zinc finger proteins  
JOURNAL Patent: US 6323335-A 3 27-NOV-2001;  
FEATURES Location/Qualifiers  
source 1..5868  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Alignment Scores: 494 Length: 5868  
Pred. No.: 43.00 Matches: 8  
Score: 88.89% Conservative: 0  
Percent Similarity: 88.89% Mismatches: 1  
Best Local Similarity: 88.89% Indels: 0  
Query Match: 86.00% Gaps: 0  
DB: 6

US-10-014-101B-40 (1-9) x AR262645 (1-5868)  
QY 1 ValLeuGlyLeuGlyGlnPheCys 9  
|||||  
Db 1728 GTTCTGGTGGCTGGGCGCTGTTCTGC 1702

RESULT 27  
AX346724  
LOCUS AX346724 6113 bp DNA linear PAT 01-FEB-2002  
DEFINITION Sequence 1795 from Patent WO0200928.  
ACCESSION AX346724  
VERSION AX346724.1 GI:18494610  
KEYWORDS synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1  
AUTHORS Olek, A.; Piepenbrock, C. and Berlin, K.  
TITLE Diagnosis of diseases associated with the immune system  
JOURNAL Patent: WO 0200928-A 1795 03-JAN-2002;  
FEATURES Epigenomics AG (DE)  
source Location/Qualifiers  
1..6113  
/organism="synthetic construct"

```

/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"

ORIGIN
Alignment Scores:
Pred. No.: 513 Length: 6113
Score: 43.00 Matches: 8
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 1
Query Match: 86.00% Indels: 0
DB: 6 Gaps: 0

US-10-014-101B-40 (1-9) x AX346724 (1-6113)

Qy 1 ValLeuGlyGlyLeuGlyGlnPheCys 9
Db 5780 GTTTGGTGGGTTAGGTTTGT 5806

RESULT 28
CQ715456/c 6177 bp DNA linear PAT 03-FEB-2004
LOCUS
DEFINITION Sequence 1390 from Patent WO02068579.
ACCESSION CQ715456
VERSION CQ715456.1 GI:42276313
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 1390 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source
Location/Qualifiers
1..6177
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 518 Length: 6177
Score: 43.00 Matches: 8
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 1
Query Match: 86.00% Indels: 0
DB: 6 Gaps: 0

US-10-014-101B-40 (1-9) x CQ715456 (1-6177)

Qy 1 ValLeuGlyGlyLeuGlyGlnPheCys 9
Db 1182 GTTCTGGTGCCCTGGCCCTGTTCTGC 1156

RESULT 29
HUMHOXY1/c 6188 bp mRNA linear PRI 10-FEB-1999
LOCUS
DEFINITION Homo sapiens mRNA for zinc-finger DNA-binding protein, complete
cDS.
D45132
VERSION D45132.1 GI:1405347
KEYWORDS heme-oxygenase-1; zinc-finger DNA-binding protein; MTB-Zf.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Murao, Y., Takahashi, K., Yoshizawa, M. and Shibahara, S.
TITLE cDNA cloning of a novel protein containing two zinc-finger domains

```

that may function as a transcription factor for the human heme-oxygenase-1 gene  
Kur. J. Biochem. 235 (3), 471-479 (1996)  
96184519  
8654390  
2 (bases 1 to 6188)  
Shibahara, S.  
Direct Submission  
Submitted (11-JAN-1995) Shigeki Shibahara, Tohoku University School  
of Medicine, Dept. of Applied Physiol. and Mol. Biol.; 2-1  
Seiryomachi, Aoba-ku, Sendai, Miyagi 980, Japan (Tel:022-717-8117,  
Fax:022-717-8118)  
On Jul 6, 1996 this sequence version replaced gi:1389628.  
Sequence updated (30-May-96) by : Shigeki Shibahara.

## FEATURES

## source

```

Location/Qualifiers
1..6188
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="THP-1"
/tissue_type="kidney"
178..4626
/codon_start=1

```

## CDS

```

/product="zinc-finger DNA-binding protein"
/protein_id="BAA08110.1"
/db_xref="GI:1405348"
/translation="MRDSAGPGDEEKPDSASALEQATLQEVASQVPELATPAPA
WEFPPERLEAAACEVDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE
SVKEPEIRDEKPEDLEETSETLEDSEVTPAMQIPRTKEANGDYFETMFPF
CQCHERKFTQGLERHMHII STVNHAFKCYGKAGFTQINRRHRRHEAGLKRK
PSOTLQSEDLAGKASNGENSDSPSPGPDCLIMSEKASQDTINSVVEENG
EVEILPKCKYCKVGTHTNMRRHRRVHERHLIPKGYRKGGLPEEPQPPAQATQ
NLYVSTPFEPEGADDDYIMDISNISENLNLYIDGKIQTNNNTSCDVIEGASA
DLYGNCILGTLPTVITQINIKITQVPTEDLPKGLSTNSEAKKRTASPPALPKIK
ABTSDMPVSCSLPLSISTTEAVSFHKEKSVLSKJLQLQTDKLTTPAGISA
TEIAKLGVCVSPAPASMLPVTSSRFKRTSPSPQSPALRDFGKPSDGKAWTDA
GLTSKGLSHSDSPAMSLGRDERETVSPCFDEYKMSKEWTASSAFSSVCNQPL
DLSSGVOKARGTGKTPVOWESVLDLSVHKHCDSDSEGEKESHVQPTCSAVKRRK
PTTCLMOKVLLNEYNGIDLVEYNADGTRSPSPCKSLEAQDPDLGPGSGPPATVES
TDFVCPSPALQTSLSGQLPLLIPTDPSPPFPVLTATVATPPLLTPLPAP
SSSSSPFPCSPSLNAQSPILISPTVSPSPSPPIPVPEPLMSAASGPTLSSSS
SSSSSPFSSSSSSSSSPPLSISVSSVSSGDNLEASLPMISFQBELENEGKPRE
EPQAAEQDVVQETFNKCNVCEGPFLLIKDLTKHLSIHABEWPFCFCVQLFK
DKTDLSEHRLHGVGNI FVCSVCKGFALCNLQHQHDLHPKVCCTHHEFEGTLR
PQNTDPSKAVHEMQSLPEPDLSTKEEELNDSSSELYTTIKIMASGIKTKGDVVR
LGLAQHYSPFKPFPQYHRRNPMGIGVTATNTFTNIPQTTTALRCTCKGKGVNMP
ELKHILACASDKKRTYPMKNPVPLKQVQKNGVVVLDNKGNAFRRWGQPKRLN
FVSLSKMSSNKLKLNALKKQNLQVKAQLQKNKSAKQADLNACSSSSHCPCYCNR
EFTYIGSLNKHAFCPCPKPLSPKKVSHSKGSHSSPASSDKNSNHRRTADA
BIRMQSMOTPLGKTRASSGPTQVPLPSSSPSKQNVKFAASVKSKPSSSLKNSSP
IRMAKITHVEGKPKAVAKNHS AOLSSKTSRSLHVRVQKSAVLQSKSTLASKKRTDR
FNKRSRSGGPTVRSIQLAAADLSENKREDGSAKQELKDFRNL"

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 519 Length: 6188
Score: 43.00 Matches: 8
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 1
Query Match: 86.00% Indels: 0
DB: 9 Gaps: 0

```

US-10-014-101B-40 (1-9) x HUMHOXY1 (1-6188)

```

Qy 1 ValLeuGlyGlyLeuGlyGlnPheCys 9
Db 1182 GTTCTGGTGCCCTGGCCCTGTTCTGC 1156

```

## RESULT 30

AX345583

LOCUS 6436 bp DNA

DEFINITION Sequence 654 from Patent WO0200928.

AX345583

AX345583

linear

PAT 01-FEB-2002

```

VERSION  AX345583.1  GI:18493469
KEYWORDS  .
SOURCE    synthetic construct
ORGANISM  synthetic construct
          other sequences; artificial sequences.
REFERENCE 1
AUTHORS   Olek, A., Piepenbrock, C. and Berlin, K.
TITLE     Diagnosis of diseases associated with the immune system
JOURNAL   Epigenomics AG (DE)
          Patent: WO 0200928-A 654 03-JAN-2002;
          Epigenomics AG (DE)
FEATURES  Location/Qualifiers
          source
            1..6436
              /organism="synthetic construct"
              /mol_type="unassigned DNA"
              /db_xref="taxon:32630"
              /note="chemically treated genomic DNA (Homo sapiens)"
ORIGIN
Alignment Scores:
Pred. No.:      538      Length:      6436
Score:          43.00    Matches:      8
Percent Similarity: 88.89%  Conservative: 0
Best Local Similarity: 88.89%  Mismatches: 1
Query Match:    86.00%    Indels:    0
DB:             6        Gaps:      0
US-10-014-101B-40 (1-9) x AX345583 (1-6436)
QY      1 ValLeuGlyGlyLeuGlyGlnPheCys 9
Db      5764 GTTTGGTGGGTAGGTGTTTTCG 5790

RESULT 31
AX780015/c
LOCUS     AX780015      6704 bp      DNA      linear      PAT 14-JUL-2003
DEFINITION Sequence 2172 from Patent WO03039443.
ACCESSION AX780015
VERSION   AX780015.1  GI:32697009
KEYWORDS  .
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS   Haferlach, T., Schoch, C., Kern, W., Kohlmann, A., Schnittger, S.,
          Dugas, M., Eils, R., Brors, B. and Mergenthaler, S.
TITLE     Novel genetic markers for leukemias
JOURNAL   Patent: WO 03039443-A 2172 15-MAY-2003;
          Deutsches Krebsforschungszentrum (DE);
          Ludwig-Maximilians-Universitaet Muenchen (DE); Haferlach, Torsten,
          PD Dr. Dr. (DE); Schoch, Claudia (DE); Kern, Wolfgang (DE)
FEATURES  Location/Qualifiers
          source
            1..6704
              /organism="Homo sapiens"
              /mol_type="unassigned DNA"
              /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.:      559      Length:      6704
Score:          43.00    Matches:      8
Percent Similarity: 88.89%  Conservative: 0
Best Local Similarity: 88.89%  Mismatches: 1
Query Match:    86.00%    Indels:    0
DB:             6        Gaps:      0
US-10-014-101B-40 (1-9) x AX780015 (1-6704)
QY      1 ValLeuGlyGlyLeuGlyGlnPheCys 9
Db      1476 GTTCTGGTGCCCTGGGCCCTGCTTCG 1450

RESULT 32
AX281287
LOCUS     AX281287      7603 bp      DNA      linear      PAT 02-NOV-2001
DEFINITION Sequence 29 from Patent WO0177164.
ACCESSION AX281287
VERSION   AX281287.1  GI:16608542
KEYWORDS  .
SOURCE    synthetic construct
ORGANISM  synthetic construct
          other sequences; artificial sequences.
REFERENCE 1
AUTHORS   Olek, A., Piepenbrock, C. and Berlin, K.
TITLE     Diagnosis of diseases associated with apoptosis
JOURNAL   Patent: WO 0177164-A 29 18-OCT-2001;
          Epigenomics AG (DE)
FEATURES  Location/Qualifiers
          source
            1..7603
              /organism="synthetic construct"
              /mol_type="unassigned DNA"
              /db_xref="taxon:32630"
              /note="chemically treated genomic DNA (Homo sapiens)"
ORIGIN
Alignment Scores:
Pred. No.:      628      Length:      7603
Score:          43.00    Matches:      8
Percent Similarity: 88.89%  Conservative: 0
Best Local Similarity: 88.89%  Mismatches: 1
Query Match:    86.00%    Indels:    0
DB:             6        Gaps:      0
US-10-014-101B-40 (1-9) x AX281287 (1-7603)
QY      1 ValLeuGlyGlyLeuGlyGlnPheCys 9
Db      334 GTTTGGTGGGTAGGTGTTTTCG 360

RESULT 33
AX345210
LOCUS     AX345210      7603 bp      DNA      linear      PAT 01-FEB-2002
DEFINITION Sequence 281 from Patent WO0200928.
ACCESSION AX345210
VERSION   AX345210.1  GI:18493096
KEYWORDS  .
SOURCE    synthetic construct
ORGANISM  synthetic construct
          other sequences; artificial sequences.
REFERENCE 1
AUTHORS   Olek, A., Piepenbrock, C. and Berlin, K.
TITLE     Diagnosis of diseases associated with the immune system
JOURNAL   Patent: WO 0200928-A 281 03-JAN-2002;
          Epigenomics AG (DE)
FEATURES  Location/Qualifiers
          source
            1..7603
              /organism="synthetic construct"
              /mol_type="unassigned DNA"
              /db_xref="taxon:32630"
              /note="chemically treated genomic DNA (Homo sapiens)"
ORIGIN
Alignment Scores:
Pred. No.:      628      Length:      7603
Score:          43.00    Matches:      8
Percent Similarity: 88.89%  Conservative: 0
Best Local Similarity: 88.89%  Mismatches: 1
Query Match:    86.00%    Indels:    0
DB:             6        Gaps:      0
US-10-014-101B-40 (1-9) x AX345210 (1-7603)
QY      1 ValLeuGlyGlyLeuGlyGlnPheCys 9
Db      334 GTTTGGTGGGTAGGTGTTTTCG 360
```

```
RESULT 34
HSU17838/c
LOCUS       HSU17838             7942 bp    mRNA    linear    PRI 31-AUG-2000
DEFINITION  Homo sapiens zinc finger protein RIZ mRNA, complete cds.
ACCESSION   U17838
VERSION     U17838.1  GI:1669774
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 7942)
AUTHORS     Bayse,I.M., Shao,G. and Huang,S.
TITLE       The retinoblastoma protein binds to RIZ, a zinc-finger protein that
            shares an epitope with the adenovirus E1A protein
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 92 (10), 4467-4471 (1995)
MEDLINE     95273384
PUBMED      7538672
REFERENCE   2  (bases 1 to 7942)
AUTHORS     Huang,S.
TITLE       Direct Submission
JOURNAL     Submitted (29-NOV-1994) Shi Huang, La Jolla Cancer Research
            Foundation, 10901 N. Torrey Pines Road, La Jolla, CA 92037, USA
REFERENCE   3  (bases 1 to 7942)
AUTHORS     Huang,S.
TITLE       Direct Submission
JOURNAL     Submitted (31-AUG-2000) Shi Huang, La Jolla Cancer Research
            Foundation, 10901 N. Torrey Pines Road, La Jolla, CA 92037, USA
REMARK      Amino acid sequence updated by submitter
COMMENT      On Nov 15, 1996 this sequence version replaced gi:949997.
FEATURES
            Location/Qualifiers
             source
              1..7942
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               856..6015
               /function="zinc finger protein"
               /codon_start=1
               /product="zinc finger protein RIZ"
               /protein_id="AAC50820.2"
               /db_xref="GI:9955379"
               /translation="MQNTEPVAATETLAEPVHVLGLPEVRLFPSPAVDKTRIGV
            WATPKLKGKGFVFGDKKRSQVKNVYMEVYVNPNGWICIDTDEKGNWLRYV
            NWACSEBQNLFFLEIRALYKLPAPGELLVYNGEDNPEIAAIEBERASAR
            SKSSPKRKKGNKQENKNGKIQIDIKTSEPDFTSANMRDSEAGKDEEKPFA
            SALEQATILQVASQVPPPELATPAPWEPQEPDERLEAAACEVNDLGEDEEKEE
            DEEEEDDDDELEDEGEDEEASMPNENSVKEPEIRCDKPEDLLPEKPTTSRETLDC
            SEVTPAMQIPRTKEBANGDVETFMFPFPCOCHERKFTTKOGLERHWHIHLSTVNHAFK
            KYGKAGFTQILARRHERHEAGLKRKPSQTLQPSDLADGKASGENVAKDDSPPS
            LGPDCLIMSEKASQDITNSVVEENGVEVKEHPCKYCKKVFGTTHMRHQRVHER
            HLIPKVRKRGLEEFPPPARQAQATQNVYVFTSEPEEGEADDVYINDISNISENL
            NYIDGKIQTNNNTSNDVIEMSASADLYGINCLLTPTVTEITQNIKTQVPTEDL
            PREPLGTSNEAKRRTASPPALPKIKAEITDSDPMVPSCLSLPLSISTEAVSFHKE
            KSVLSKQLLODOKLTTPAGISATRIAKLGPVCVSAPASMLPVTSSRFKRTSS
            PSSSPKALLRDPFGKDGKAWTDAGLTSSKSLKESDSPAFLSGLRDRFVSP
            PCFDEYMKSEWTASAFSVNCNQQLDLSGKVKAGETKIPQWVESLVLDVHKK
            HCSDEGKFKESHVQPTCSAVKRRKPTCMQLKVLNENGLDLPVENPADGTRSP
            SPCKLEAQPDPGFPAPFVETSDVCFSPFALQTPSLSGQPLPLIPTDPS
            SPPCPFVLTPATPPPLTPVLPAPSSASPHPCPSPLSNATQSPILSPVSP
            SPPIPPVPLPMSAAGSPPTLSSSSSSSSSSSSSSSSSSSSSPPLSAISVSVSS
            GDNLEASLPMISFKQELNENGLKPEEPQSAEODVVQETFNKFNVCNCESPFLS
            IKDQLHSLHAEVWFKEFCFQVLPKDKIDLSEHFLHGVGNIFVCSVKCKEFAFL
            CNLQQRDLSHDEQNTHEFSGTLRQNFDTDPSKAHVHMQSLPEDPLETSKEEE
            LNDSSBELYTTIKAMAGTIKDPDVLRLGNQHFSPKPPFYHHRNPMGIGVATN
            FTTHNIPOFTTKTIRAKTGKGVNDNMPHLKHIACASADKKRVTPKNVPVLPKQTV
            QPKNGVLDNSGNKARFMQGPKRNFSEVLSKMSNKLKNAKLNKLVOKAILQ
            KNKSAKQADLNACSESSHICPYCNREFTYIGLNKHAFFSCPKPLSPKPKYVSHS
            SKKGHS SPASSDANSNHRRTADAEIKMQSMQTPICKTRARSGTQVLPSSSF
            RSKQNVRAASVKAPSSSLRNSPPIRMAKITHVEGKKPAKVNKNHSAQISKTSR
            SLHVRVQSKAVQKSLASKKRTDRFNKISRERSGEPVTSGLAAADASNKRE
            DGSAKQLDFSYSLASRCSPPAPYITTOYRKVKYKAPAAAFQFQFPFFKE"
```

```
Alignment Scores:
Pred. No.: 654          Length: 7942
Score: 43.00           Matches: 8
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 1
Query Match: 86.00% Indels: 0
DB: 9 Gaps: 0

US-10-014-101B-40 (1-9) x HSU17838 (1-7942)
QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9
Db 2463 GTTCTGGTGGCTGGCGCTGTCTGC 2437

RESULT 35
LOCUS       AX323793             9504 bp    DNA    linear    PAT 07-JAN-2002
DEFINITION  Sequence 281 from Patent WO0192565.
ACCESSION   AX323793
VERSION     AX323793.1  GI:18094541
KEYWORDS    synthetic construct
SOURCE      synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
AUTHORS     Olek,A., Piepenbrock,C. and Berlin,K.
TITLE       Diagnosis of diseases associated with dna transcription
JOURNAL     Patent: WO 0192565-A 281 06-DEC-2001;
            Epigenomics AG (DE)
FEATURES
            Location/Qualifiers
             source
              1..9504
               /organism="synthetic construct"
               /mol_type="unassigned DNA"
               /db_xref="taxon:32630"
               /note="Chemically treated genomic DNA (Homo sapiens)"

ORIGIN

Alignment Scores:
Pred. No.: 771          Length: 9504
Score: 43.00           Matches: 8
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 1
Query Match: 86.00% Indels: 0
DB: 6 Gaps: 0

US-10-014-101B-40 (1-9) x AX323793 (1-9504)
QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9
Db 6626 GTTTGGTGGTGGTTAGGTGTTTGT 6652

RESULT 36
AC023042/c
LOCUS       AC023042             55622 bp    DNA    linear    HTG 30-JUL-2002
DEFINITION  Homo sapiens clone RP11-141M19, LOW-PASS SEQUENCE SAMPLING.
ACCESSION   AC023042
VERSION     AC023042.3  GI:22004286
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 55622)
AUTHORS     Birren,B., Nusbaum,C. and Lander,E.
TITLE       Homo sapiens, clone RP11-141M19
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 55622)
AUTHORS     Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Bada,F.,
            Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            Dearellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
            Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
```

Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
 Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,  
 Landers, T., Lehotsky, J., Levine, R., Lieu, C., Liu, G., Locke, K.,  
 Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,  
 McPheeters, R., Meldrim, J., Meneus, L., Morrow, J., Naylor, J.,  
 Norman, C. H., O'Connor, T., O'Donnell, P., Oliver, T. M., Peterson, K.,  
 Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,  
 Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,  
 Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
 Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,  
 Zimmer, A. and Zody, M.

# TITLE JOURNAL

Submitted (07-FEB-2000) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 55622)

# REFERENCE AUTHORS

Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,  
 Barrna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B.,  
 Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,  
 Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,  
 Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,  
 Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,  
 Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,  
 Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,  
 Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,  
 McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,  
 Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H.,  
 O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,  
 Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,  
 Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P.,  
 Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,  
 Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,  
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,  
 Zembek, L., Zimmer, A. and Zody, M.

# TITLE JOURNAL

Submitted (30-JUL-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jul 30, 2002 this sequence version replaced gi:7263244.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L941

Center clone name: 141\_M\_19

-----

\* NOTE: This record contains 71 individual

\* sequencing reads that have not been assembled into

\* contigs. Runs of N are used to separate the reads

\* and the order in which they appear is completely

\* arbitrary. Low-pass sequence sampling is useful for

\* identifying clones that may be gene-rich and allows

\* overlap relationships among clones to be deduced.

\* However, it should not be assumed that this clone

\* will be sequenced to completion. In the event that

\* the record is updated, the accession number will

\* be preserved.

\* 1 665: contig of 665 bp in length

\* 666 765: gap of 100 bp

\* 766 1446: contig of 681 bp in length

\* 1447 1546: gap of 100 bp

\* 1547 2243: contig of 697 bp in length

\* 2244 2343: gap of 100 bp

\* 2344 3036: contig of 693 bp in length

\* 3037 3136: gap of 100 bp

\* 3137 3805: contig of 669 bp in length

\* 3806 3905: gap of 100 bp

\* 3906 4586: contig of 681 bp in length

\* 4587 4686: gap of 100 bp

\* 4687 5348: contig of 662 bp in length

5349 5448: gap of 100 bp

5449 5441: contig of 693 bp in length

6142 6241: gap of 100 bp

6242 6931: contig of 690 bp in length

6932 7031: gap of 100 bp

7032 7723: contig of 692 bp in length

7724 7823: gap of 100 bp

7824 8519: contig of 696 bp in length

8520 8619: gap of 100 bp

8620 9315: contig of 696 bp in length

9316 9415: gap of 100 bp

9416 10116: contig of 701 bp in length

10117 10216: gap of 100 bp

10217 10890: contig of 674 bp in length

10891 10990: gap of 100 bp

10991 11666: contig of 676 bp in length

11667 11766: gap of 100 bp

11767 12437: contig of 671 bp in length

12438 12537: gap of 100 bp

12538 12530: contig of 693 bp in length

12539 13330: gap of 100 bp

13331 14016: contig of 686 bp in length

14017 14116: gap of 100 bp

14117 14802: contig of 686 bp in length

14803 14902: gap of 100 bp

14904 15365: contig of 663 bp in length

15366 15665: gap of 100 bp

15666 16328: contig of 663 bp in length

16329 16428: gap of 100 bp

16429 17106: contig of 678 bp in length

17107 17206: gap of 100 bp

17207 17911: contig of 705 bp in length

17912 18011: gap of 100 bp

18012 18697: contig of 686 bp in length

18698 18797: gap of 100 bp

18798 19491: contig of 694 bp in length

19492 19591: gap of 100 bp

19592 20264: contig of 673 bp in length

20265 20364: gap of 100 bp

20365 21055: contig of 691 bp in length

21056 21155: gap of 100 bp

21156 21850: contig of 695 bp in length

21851 21950: gap of 100 bp

21951 22634: contig of 684 bp in length

22635 22734: gap of 100 bp

22735 23389: contig of 655 bp in length

23390 23489: gap of 100 bp

23490 24172: contig of 683 bp in length

24173 24272: gap of 100 bp

24273 24962: contig of 690 bp in length

24963 25062: gap of 100 bp

25063 25751: contig of 689 bp in length

25752 25851: gap of 100 bp

25852 26523: contig of 672 bp in length

26524 26523: gap of 100 bp

26524 27320: contig of 697 bp in length

27321 27420: gap of 100 bp

27421 28086: contig of 666 bp in length

28087 28186: gap of 100 bp

28187 28968: contig of 682 bp in length

28969 28968: gap of 100 bp

28969 29649: contig of 681 bp in length

29650 29749: gap of 100 bp

29750 30440: contig of 691 bp in length

30441 30540: gap of 100 bp

30541 31230: contig of 690 bp in length

31231 31330: gap of 100 bp

31331 32013: contig of 683 bp in length

32014 32113: gap of 100 bp

32114 32800: contig of 687 bp in length

32801 32900: gap of 100 bp

32901 33590: contig of 690 bp in length

33591 33690: gap of 100 bp

```

* 33691 34385: contig of 695 bp in length
* 34386 34485: gap of 100 bp
* 34486 35178: contig of 693 bp in length
* 35179 35278: gap of 100 bp
* 35279 35964: contig of 686 bp in length
* 35965 36064: gap of 100 bp
* 36065 36763: contig of 699 bp in length
* 36764 36863: gap of 100 bp
* 36864 37522: contig of 659 bp in length
* 37523 37622: gap of 100 bp
* 37623 38299: contig of 673 bp in length
* 38299 38395: gap of 100 bp
* 38396 39028: contig of 633 bp in length
* 39029 39128: gap of 100 bp
* 39129 39810: contig of 682 bp in length
* 39811 39910: gap of 100 bp
* 39911 40581: contig of 671 bp in length
* 40582 40681: gap of 100 bp
* 40682 41363: contig of 682 bp in length
* 41364 41463: gap of 100 bp
* 41464 42151: contig of 687 bp in length
* 42152 42250: gap of 100 bp
* 42251 42921: contig of 671 bp in length
* 42922 43021: gap of 100 bp
* 43022 43705: contig of 684 bp in length
* 43706 43805: gap of 100 bp
* 43806 44493: contig of 688 bp in length
* 44494 44593: gap of 100 bp
* 44594 45285: contig of 692 bp in length
* 45286 45385: gap of 100 bp
* 45386 46072: contig of 687 bp in length
* 46073 46172: gap of 100 bp
* 46173 46864: contig of 692 bp in length
* 46865 46964: gap of 100 bp

```

## Alignment Scores:

```

Pred. No.: 3,94e+03 Length: 55622
Score: 43.00 Matches: 7
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 87.50% Mismatches: 0
Query Match: 86.00% Indels: 0
DB: 2 Gaps: 0

```

US-10-014-101B-40 (1-9) x AC023042 (1-55622)

QY 2 LeuGlyClyLeuGlyGlnPheCys 9

Db 43326 CTGGAGGCTTGGAGAAATTTC 43303

## RESULT 37

```

AC091556
LOCUS Homo sapiens chromosome 12 clone RP11-40I110 map 12, LOW-PASS
DEFINITION SEQUENCE SAMPLING.
AC091556

```

AC091556 AC091556.1 GI:13959161

VERSION HTG; HTGS PHASE0.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 59433)

AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.

TITLE Homo sapiens chromosome 12, clone RP11-40I110

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 59433)

AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,

Barna, N., Bastien, V., Boguslavskiy, L., Bouckgalter, B., Brown, A.,

Camarata, J., Campopiano, A., Chang, J., Choepel, Y., Colangelo, M.,

Collins, J., Collymore, A., Cooke, P., DeArellano, K., Dewar, K.,

Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D.,

Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,

Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W.,

Iliev, I., Johnson, R., Jones, C., Karatas, A., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Marquis, N., Matthews, C., McCarthy, M., McSwan, P., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Sougne, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

## Direct Submission

Submitted (07-MAY-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 59433)

## REFERENCE

## AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Bouckgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Choepel, Y., Colangelo, M., Collins, J., Collymore, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karatas, A., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Sougne, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

## Direct Submission

Submitted (28-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information

Center project name: L11612

Center clone name: 401\_I\_10

-----

\* NOTE: This record contains 71 individual  
 \* sequencing reads that have not been assembled into  
 \* contigs. Runs of N are used to separate the reads  
 \* and the order in which they appear is completely  
 \* arbitrary. Low-pass sequence sampling is useful for  
 \* identifying clones that may be gene-rich and allows  
 \* overlap relationships among clones to be deduced.  
 \* However, it should not be assumed that this clone  
 \* will be sequenced to completion. In the event that  
 \* the record is updated, the accession number will  
 \* be preserved.

```

* 1 740: contig of 740 bp in length
* 741 840: gap of 100 bp
* 841 1601: contig of 761 bp in length
* 1602 1701: gap of 100 bp
* 1702 2413: contig of 712 bp in length
* 2414 2513: gap of 100 bp
* 2514 3246: contig of 733 bp in length
* 3247 3346: gap of 100 bp

```

3347	4060:	contig	of 714	bp	in length
4061	4160:	gap	of 100	bp	in length
4161	4911:	contig	of 751	bp	in length
4912	5011:	gap	of 100	bp	in length
5011	5767:	contig	of 756	bp	in length
5768	5867:	gap	of 100	bp	in length
5868	6623:	contig	of 756	bp	in length
6624	6723:	gap	of 100	bp	in length
6724	7481:	contig	of 758	bp	in length
7482	7581:	gap	of 100	bp	in length
7582	8324:	contig	of 743	bp	in length
8325	8424:	gap	of 100	bp	in length
8425	9122:	contig	of 698	bp	in length
9123	9222:	gap	of 100	bp	in length
9223	9965:	contig	of 743	bp	in length
9966	10065:	gap	of 100	bp	in length
10066	10773:	contig	of 708	bp	in length
10774	10873:	gap	of 100	bp	in length
10874	11595:	contig	of 722	bp	in length
11596	11695:	gap	of 100	bp	in length
11696	12433:	contig	of 738	bp	in length
12434	12533:	gap	of 100	bp	in length
12534	13270:	contig	of 737	bp	in length
13271	13370:	gap	of 100	bp	in length
13371	14120:	contig	of 750	bp	in length
14121	14220:	gap	of 100	bp	in length
14221	14946:	contig	of 726	bp	in length
14947	15046:	gap	of 100	bp	in length
15047	15811:	contig	of 765	bp	in length
15812	15911:	gap	of 100	bp	in length
15912	16640:	contig	of 729	bp	in length
16641	16740:	gap	of 100	bp	in length
16741	17439:	contig	of 699	bp	in length
17440	17539:	gap	of 100	bp	in length
17540	18271:	contig	of 732	bp	in length
18272	18371:	gap	of 100	bp	in length
18372	19133:	contig	of 762	bp	in length
19134	19233:	gap	of 100	bp	in length
19234	19986:	contig	of 753	bp	in length
19987	20086:	gap	of 100	bp	in length
20087	20849:	contig	of 763	bp	in length
20850	20949:	gap	of 100	bp	in length
20950	21709:	contig	of 760	bp	in length
21710	21809:	gap	of 100	bp	in length
21810	22563:	contig	of 754	bp	in length
22564	22663:	gap	of 100	bp	in length
22664	23381:	contig	of 718	bp	in length
23382	23481:	gap	of 100	bp	in length
23482	24223:	contig	of 748	bp	in length
24223	24329:	gap	of 100	bp	in length
24330	25067:	contig	of 738	bp	in length
24330	25167:	gap	of 100	bp	in length
25068	25896:	contig	of 729	bp	in length
25168	25996:	gap	of 100	bp	in length
25997	26754:	contig	of 758	bp	in length
26755	26854:	gap	of 100	bp	in length
26855	27586:	contig	of 732	bp	in length
27587	27686:	gap	of 100	bp	in length
27687	28427:	contig	of 741	bp	in length
28428	28527:	gap	of 100	bp	in length
28528	29292:	contig	of 765	bp	in length
29293	29392:	gap	of 100	bp	in length
29393	30153:	contig	of 761	bp	in length
30154	30253:	gap	of 100	bp	in length
30254	31004:	contig	of 751	bp	in length
31005	31104:	gap	of 100	bp	in length
31105	31858:	contig	of 754	bp	in length
31859	31958:	gap	of 100	bp	in length
31959	32700:	contig	of 742	bp	in length
32701	32800:	gap	of 100	bp	in length
32801	33526:	contig	of 726	bp	in length
33527	33626:	gap	of 100	bp	in length
33627	34396:	contig	of 770	bp	in length

*	34397	34496:	gap of 100 bp
*	34497	35244:	contig of 748 bp in length
*	35245	35344:	gap of 100 bp
*	35345	36086:	contig of 742 bp in length
*	36087	36186:	gap of 100 bp
*	36187	36956:	contig of 770 bp in length
*	36957	37056:	gap of 100 bp
*	37057	37817:	contig of 761 bp in length
*	37818	37917:	gap of 100 bp
*	37918	38658:	contig of 741 bp in length
*	38659	38758:	gap of 100 bp
*	38759	39505:	contig of 747 bp in length
*	39506	39605:	gap of 100 bp
*	39606	40342:	contig of 737 bp in length
*	40343	40442:	gap of 100 bp
*	40443	41154:	contig of 712 bp in length
*	41155	41254:	gap of 100 bp
*	41255	41839:	contig of 585 bp in length
*	41840	41939:	gap of 100 bp
*	41940	42679:	contig of 740 bp in length
*	42680	42779:	gap of 100 bp
*	42780	43510:	contig of 731 bp in length
*	43511	43610:	gap of 100 bp
*	43611	44361:	contig of 751 bp in length
*	44362	44461:	gap of 100 bp
*	44462	45222:	contig of 761 bp in length
*	45223	45322:	gap of 100 bp
*	45323	46061:	contig of 739 bp in length
*	46062	46161:	gap of 100 bp
*	46162	46901:	contig of 740 bp in length
*	46902	47001:	gap of 100 bp
*	47002	47764:	contig of 763 bp in length

**Alignment Scores:**

Alignment Scores:	
Pred. No.:	4.19e+03
Score:	43.00
Percent Similarity:	100.00%
Best Local Similarity:	87.50%
Query Match:	86.00%
DB:	2
Gaps:	0
Indels:	0
Mismatches:	0
Conservative:	1
Matches:	7
Length:	59433

US-10-014-101B-40 (1-9) x AC091556 (1-59433)

2 LeuGlyGlyLeuGlyGlnpheCys 9

**QY**            2   LeuGIyGlyLeuGIyGlnFneCyS 3  
||| ||| ||| ||| ||| ||| : :: |||  
**D6**            18622 CTGGAGGTCTTGGAGAA TTTTGC 18645

PESTIC. 38

RESULT 38  
AC011925

AC011925  
LOCUS AC011925 76891 bp DNA linear HTG 13-JUL-2000  
DEFINITION Homo sapiens clone RP11-15M20. LOW-PASS SEQUENCE SAMPLING.

DEFINITION HOMIO PAP  
ACCESSION AC011925

ACCESSION	AC011925	
VERSION	AC011925.2	GT:7144915

VERSION ACUL1923.2 GI:71  
KEYWORDS HTG: HTGS PHASE0

KEYWORDS  
SOURCE  
HIS, HIS\_FIRSTO:  
Homo sapiens (human)

**SOURCE :**  
**ORGANISM**

ORGANISM

REFERENCE 1 (bases 1 to 76891)  
Mammalia, Eucelia, Elmidae, Cucullariinae, Mon-

REFERENCE  
AUTHORS  
I (BASES 1 CO 76651)  
Birren, B., Linton, L., Nusbaum, C. and Lander, E.

AUTHORS	TITLE
ATTEN, B.:	Homo sapien

FREE JOURNAL Unpublished Manuscripts, Comments

2-7 (bases 1 to 76991)

REFERENCE  
AUTHORS

Barren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Baldwin, J., Barna, N., Beckerly, R., Boguski, L., Bouckgalter, B., Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M., Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Lehotsky, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,

Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,  
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
 Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,  
 Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

# TITLE JOURNAL

Submitted (16-OCT-1999) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Mar 3, 2000 this sequence version replaced gi:6056258.  
 All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L3490

Center clone name: 15\_M\_20

-----

\* NOTE: This record contains 81 individual  
 \* sequencing reads that have not been assembled into  
 \* contigs. Runs of N are used to separate the reads  
 \* and the order in which they appear is completely  
 \* arbitrary. Low-pass sequence sampling is useful for  
 \* identifying clones that may be gene-rich and allows  
 \* overlap relationships among clones to be deduced.  
 \* However, it should not be assumed that this clone  
 \* will be sequenced to completion. In the event that  
 \* the record is updated, the accession number will  
 \* be preserved.

1  
 876: contig of 876 bp in length  
 877 976: gap of 100 bp  
 1877: contig of 901 bp in length  
 1878 1977: gap of 100 bp  
 1978 2831: contig of 854 bp in length  
 2832 2931: gap of 100 bp  
 2932 3752: contig of 821 bp in length  
 3753 3852: gap of 100 bp  
 3853 4693: contig of 841 bp in length  
 4694 5632: contig of 839 bp in length  
 4794 5633: gap of 100 bp  
 5633 5732: contig of 854 bp in length  
 5733 6586: gap of 100 bp  
 6587 7526: contig of 840 bp in length  
 7527 7626: gap of 100 bp  
 7627 8483: contig of 857 bp in length  
 8484 8583: gap of 100 bp  
 8584 9424: contig of 841 bp in length  
 9425 9524: gap of 100 bp  
 9525 10382: contig of 858 bp in length  
 10383 10482: gap of 100 bp  
 10483 11355: contig of 873 bp in length  
 11356 11455: gap of 100 bp  
 11456 12330: contig of 875 bp in length  
 12331 12430: gap of 100 bp  
 12431 13230: contig of 800 bp in length  
 13231 13330: gap of 100 bp  
 13331 14172: contig of 842 bp in length  
 14173 14272: gap of 100 bp  
 14273 15104: contig of 832 bp in length  
 15105 15204: gap of 100 bp  
 15205 16034: contig of 830 bp in length  
 16035 16134: gap of 100 bp  
 16135 16945: contig of 811 bp in length  
 16946 17045: gap of 100 bp  
 17046 17860: contig of 815 bp in length  
 17861 17960: gap of 100 bp  
 17961 18842: contig of 882 bp in length  
 18843 18942: gap of 100 bp  
 18943 19801: contig of 859 bp in length  
 19802 19901: gap of 100 bp

19902 20749: contig of 848 bp in length  
 20849: gap of 100 bp  
 20850 21703: contig of 854 bp in length  
 21704 21803: gap of 100 bp  
 21804 22671: contig of 868 bp in length  
 22672 22771: gap of 100 bp  
 22772 23613: contig of 842 bp in length  
 23614 23713: gap of 100 bp  
 23714 24574: contig of 861 bp in length  
 24575 24674: gap of 100 bp  
 24675 25527: contig of 853 bp in length  
 25528 25627: gap of 100 bp  
 25628 26478: contig of 851 bp in length  
 26479 26578: gap of 100 bp  
 26579 27429: contig of 851 bp in length  
 27430 27529: gap of 100 bp  
 27530 28382: contig of 853 bp in length  
 28383 28482: gap of 100 bp  
 28483 29360: contig of 878 bp in length  
 29361 29460: gap of 100 bp  
 29461 30284: contig of 824 bp in length  
 30285 30384: gap of 100 bp  
 30385 31252: contig of 868 bp in length  
 31253 31352: gap of 100 bp  
 31353 32218: contig of 866 bp in length  
 32219 32318: gap of 100 bp  
 32319 33150: contig of 832 bp in length  
 33151 33250: gap of 100 bp  
 33251 34083: contig of 833 bp in length  
 34084 34183: gap of 100 bp  
 34184 34971: contig of 788 bp in length  
 34972 35071: gap of 100 bp  
 35072 35896: contig of 825 bp in length  
 35897 35996: gap of 100 bp  
 35997 36872: contig of 876 bp in length  
 36873 36972: gap of 100 bp  
 36973 37855: contig of 883 bp in length  
 37856 37955: gap of 100 bp  
 37956 38789: contig of 834 bp in length  
 38790 38889: gap of 100 bp  
 38890 39747: contig of 858 bp in length  
 39748 39847: gap of 100 bp  
 39848 40719: contig of 872 bp in length  
 40720 40819: gap of 100 bp  
 40820 41690: contig of 871 bp in length  
 41691 41790: gap of 100 bp  
 41791 42639: contig of 849 bp in length  
 42640 42739: gap of 100 bp  
 42740 43581: contig of 842 bp in length  
 43582 43681: gap of 100 bp  
 43682 44476: contig of 795 bp in length  
 44477 44576: gap of 100 bp  
 44577 45396: contig of 820 bp in length  
 45397 45496: gap of 100 bp  
 45497 46329: contig of 833 bp in length  
 46330 46429: gap of 100 bp  
 46430 47294: contig of 865 bp in length  
 47295 47394: gap of 100 bp  
 47395 48258: contig of 864 bp in length  
 48259 48358: gap of 100 bp  
 48359 49197: contig of 839 bp in length  
 49198 49297: gap of 100 bp  
 49298 50178: contig of 881 bp in length  
 50179 50278: gap of 100 bp  
 50279 51141: contig of 863 bp in length  
 51142 51241: gap of 100 bp  
 51242 52114: contig of 873 bp in length  
 52115 52214: gap of 100 bp  
 52215 53049: contig of 835 bp in length  
 53050 53149: gap of 100 bp  
 53150 53998: contig of 849 bp in length  
 53999 54098: gap of 100 bp  
 54099 54934: contig of 836 bp in length

```

* 54935 55034: gap of 100 bp
* 55035 55877: contig of 843 bp in length
* 55878 55977: gap of 100 bp
* 55978 56845: contig of 868 bp in length
* 56846 56945: gap of 100 bp
* 56946 57814: contig of 869 bp in length
* 57815 57914: gap of 100 bp
* 57915 58758: contig of 844 bp in length
* 58759 58858: gap of 100 bp
* 58859 59721: contig of 863 bp in length
* 59722 59821: gap of 100 bp
* 59822 60707: contig of 886 bp in length
* 60708 60807: gap of 100 bp
* 60808 61626: contig of 819 bp in length
* 61627 61726: gap of 100 bp
* 61727 62587: contig of 861 bp in length
* 62588 62687: gap of 100 bp
* 62688 63531: contig of 844 bp in length
* 63532 63631: gap of 100 bp
* 63632 64470: contig of 839 bp in length
* 64471 64570: gap of 100 bp
* 64571 65403: contig of 833 bp in length
* 65404 66374: gap of 100 bp
* 66375 66474: gap of 100 bp
* 66475 67375: contig of 901 bp in length
* 67376 67475: gap of 100 bp
* 67476 68349: contig of 874 bp in length
* 68350 68449: gap of 100 bp
* 68450 69305: contig of 856 bp in length

```

## Alignment Scores:

```

Pred. No.: 5.31e+03 Length: 76891
Score: 43.00 Matches: 7
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 87.50% Mismatches: 0
Query Match: 86.00% Indels: 0
DB: 2 Gaps: 0

```

US-10-014-101B-40 (1-9) x AC011925 (1-76891)

Qy 2 LeuGlyGlyLeuGlyGlnPheCys 9

Db 13808 TTGGTGTCTTGGCAACTATCG 13831

## RESULT 39

AC020850\_3

WPCOMMENT

Sequence split into 5 fragments LOCUS AC020850 Accession AC020850

```

Fragment Name Begin End
AC020850_0 1 110000
AC020850_1 100001 210000
AC020850_2 200001 310000
AC020850_3 300001 410000
AC020850_4 400001 444882
Continuation (4 of 5) of AC020850 from base 300001 (AC020850 Mus musculus clone RP21-460)

```

## Alignment Scores:

```

Pred. No.: 7.39e+03 Length: 110000
Score: 43.00 Matches: 7
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 0
Query Match: 86.00% Indels: 0
DB: 2 Gaps: 0

```

US-10-014-101B-40 (1-9) x AC020850\_3 (1-110000)

Qy 1 ValLeuGlyGlyLeuGlyGlnPheCys 9

Db 87394 CTTCTGGGAGGCTTAGGCAAGTTTGT 87420

## RESULT 40

AC024999

LOCUS AC024999 132446 bp DNA linear HTG 13-JUL-2000  
DEFINITION Homo sapiens chromosome 1 clone RP11-178K15 map 1, LOW-PASS  
SEQUENCE SAMPLING.

AC024999

AC024999.2 GI:7767841

HTG: HTGS PHASE0.

Source Homo sapiens (human)

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 132446)

AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.

TITLE Homo sapiens chromosome 1, clone RP11-178K15

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 132446)

AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,  
Boguslavsky, L., Bouckgalter, B., Brown, A., Burkett, G.,  
Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,  
Collamore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,  
Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,  
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,  
Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L.,  
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,  
Klein, J., LaRocque, K., Lamazares, R., Landers, T., Lehotzky, J.,  
Levine, R., Liu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,  
McCarthy, M., McSwan, P., McGurk, A., McKernan, K., McSheeters, R.,  
Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J.,  
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,  
O'Neil, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,  
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
Roy, A., Santos, R., Schauer, S., Severy, F., Spencer, B.,  
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
Tessaye, S., Theodore, J., Tirrell, A., Travers, M., Trifilio, J.,  
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,  
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission

Submitted (03-MAR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On May 10, 2000 this sequence version replaced gi:7145043.

COMMENT All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information

Center project name: L7587

Center clone name: 178\_K15

-----

\* NOTE: This record contains 146 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

```

* 1 757: contig of 757 bp in length
* 758 857: gap of 100 bp
* 858 1609: contig of 752 bp in length
* 1610 1709: gap of 100 bp
* 1710 2474: contig of 765 bp in length
* 2475 3319: contig of 745 bp in length
* 3320 3419: gap of 100 bp
* 3420 4160: contig of 741 bp in length
* 4161 4260: gap of 100 bp

```

```
* 4261 5007: contig of 747 bp in length
* 5008 5107: gap of 100 bp
* 5108 5850: contig of 743 bp in length
* 5851 5950: gap of 100 bp
* 5951 6690: contig of 740 bp in length
* 6691 6790: gap of 100 bp
* 6791 7540: contig of 750 bp in length
* 7541 7640: gap of 100 bp
* 7641 8365: contig of 725 bp in length
* 8366 8465: gap of 100 bp
* 8466 9236: contig of 771 bp in length
* 9237 9336: gap of 100 bp
* 9337 10088: contig of 752 bp in length
* 10089 10188: gap of 100 bp
* 10189 10928: contig of 740 bp in length
* 10929 11028: gap of 100 bp
* 11029 11780: contig of 752 bp in length
* 11781 11880: gap of 100 bp
* 11881 12632: contig of 752 bp in length
* 12633 12732: gap of 100 bp
* 12733 13476: contig of 744 bp in length
* 13477 13576: gap of 100 bp
* 13577 14335: contig of 759 bp in length
* 14336 14435: gap of 100 bp
* 14436 15194: contig of 759 bp in length
* 15195 15294: gap of 100 bp
* 15295 16065: contig of 771 bp in length
* 16066 16165: gap of 100 bp
* 16166 16924: contig of 759 bp in length
* 16925 17024: gap of 100 bp
* 17025 17783: contig of 759 bp in length
* 17784 17883: gap of 100 bp
* 17884 18636: contig of 753 bp in length
* 18637 18736: gap of 100 bp
* 18737 19485: contig of 749 bp in length
* 19486 19585: gap of 100 bp
* 19586 20342: contig of 757 bp in length
* 20343 20442: gap of 100 bp
* 20443 21209: contig of 767 bp in length
* 21210 21309: gap of 100 bp
* 21310 22086: contig of 777 bp in length
* 22087 22186: gap of 100 bp
* 22187 22950: contig of 764 bp in length
* 22951 23050: gap of 100 bp
* 23051 23802: contig of 752 bp in length
* 23803 23902: gap of 100 bp
* 23903 24649: contig of 747 bp in length
* 24650 24750: gap of 100 bp
* 24750 25507: contig of 758 bp in length
* 25508 25607: gap of 100 bp
* 25608 26364: contig of 757 bp in length
* 26365 26464: gap of 100 bp
* 26465 27221: contig of 757 bp in length
* 27222 27321: gap of 100 bp
* 27322 28089: contig of 768 bp in length
* 28090 28189: gap of 100 bp
* 28190 28968: contig of 779 bp in length
* 28969 29068: gap of 100 bp
* 29069 29837: contig of 769 bp in length
* 29838 29937: gap of 100 bp
* 29938 30679: contig of 742 bp in length
* 30680 30779: gap of 100 bp
* 30780 31511: contig of 732 bp in length
* 31512 31611: gap of 100 bp
* 31612 32390: contig of 779 bp in length
* 32391 32490: gap of 100 bp
* 32491 33251: contig of 761 bp in length
* 33252 33351: gap of 100 bp
* 33352 34098: contig of 747 bp in length
* 34099 34198: gap of 100 bp
* 34199 34942: contig of 744 bp in length
* 34943 35042: gap of 100 bp
* 35043 35785: contig of 743 bp in length

* 35786 35885: gap of 100 bp
* 35886 36646: contig of 761 bp in length
* 36647 36746: gap of 100 bp
* 36747 37500: contig of 754 bp in length
* 37501 37600: gap of 100 bp
* 37601 38371: contig of 771 bp in length
* 38372 38471: gap of 100 bp
* 38472 39240: contig of 769 bp in length
* 39241 39340: gap of 100 bp
* 39341 40097: contig of 757 bp in length
* 40098 40197: gap of 100 bp
* 40198 40919: contig of 722 bp in length
* 40920 41019: gap of 100 bp
* 41020 41767: contig of 748 bp in length
* 41768 41867: gap of 100 bp
* 41868 42623: contig of 756 bp in length
* 42624 42724: gap of 100 bp
* 42724 43494: contig of 770 bp in length
* 43494 43594: gap of 100 bp
* 43594 44345: contig of 752 bp in length
* 44346 44445: gap of 100 bp
* 44446 45198: contig of 753 bp in length
* 45199 45298: gap of 100 bp
* 45299 46043: contig of 745 bp in length
* 46044 46144: gap of 100 bp
* 46144 46914: contig of 771 bp in length
* 46915 47015: gap of 100 bp
* 47015 47789: contig of 775 bp in length
* 47790 47889: gap of 100 bp
* 47890 48652: contig of 763 bp in length
* 48653 48752: gap of 100 bp
* 48753 49504: contig of 752 bp in length
* 49505 49604: gap of 100 bp
* 49605 50358: contig of 754 bp in length
* 50359 50458: gap of 100 bp
* 50459 51226: contig of 768 bp in length
* 51227 51326: gap of 100 bp
* 51327 52087: contig of 761 bp in length
* 52088 52187: gap of 100 bp
* 52188 52948: contig of 761 bp in length
* 52949 53048: gap of 100 bp
* 53049 53824: contig of 776 bp in length
* 53825 53924: gap of 100 bp
* 53925 54755: contig of 831 bp in length
* 54756 54855: gap of 100 bp
* 54856 55699: contig of 844 bp in length
* 55700 55799: gap of 100 bp
* 55800 56678: contig of 879 bp in length
* 56679 56778: gap of 100 bp
* 56779 57612: contig of 834 bp in length
* 57613 57712: gap of 100 bp
* 57713 58521: contig of 809 bp in length
* 58522 58621: gap of 100 bp
* 58622 59446: contig of 825 bp in length
```

## Alignment Scores:

Pred. No.:	8.77e+03	Length:	132446
Score:	43.00	Matches:	8
Percent Similarity:	88.89%	Conservative:	0
Best Local Similarity:	88.89%	Mismatches:	1
Query Match:	86.00%	Indels:	0
DB:	2	Gaps:	0

US-10-014-101B-40 (1-9) x AC024999 (1-132446)

Qy	1	ValLeuGlyLeuGlyClnPheCys	9
Db	51454	GTCTGGGTGGCTGGCCTGTCTGC	51480

Search completed: February 18, 2005, 05:26:00  
Job time : 1503.08 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	45	90.0	65047	11	ACN44020	ACN44020	Mouse gen
C 2	44	88.0	1509	2	AAQ64510	AAQ64510	Thermus t
3	43	86.0	503	12	ACH75978	ACH75978	Human gen
4	43	86.0	789	12	ACH89659	ACH89659	Human gen
C 5	43	86.0	5849	6	ABK17335	ABK17335	Human Rb-

# Recombinant nucleic acid useful for diagnosis and treatment of carcinoma

PT comprises a nucleotide sequence.

PS Claim 1; SEQ ID NO 259; Opp; English.

XX

CC The present invention relates to novel DNA and protein sequences which

CC are associated with carcinomas. The sequences are useful for: (i) for

CC screening drug candidates; (ii) for screening of bioactive agent capable

CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of

CC a bioactive agent capable of modulating the activity of CAP; (iv) for

CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing

CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating

CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;

CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for

CC determining Carcinoma Associated (CA) gene copy number. In addition, the

CC CA genes are useful as DNA vaccines and the CAP are useful as markers of

CC carcinoma including lymphoma. The present sequence is one such CA coding

CC sequence. Note: This patent is an equivalent to basic patent

CC US2002182586A1, for which no sequence data was published

XX

SQ Sequence 65047 BP; 15178 A; 14282 C; 14920 G; 17577 T; 0 U; 3090 Other;

Alignment Scores:

Pred. No.:	7.53e+03	Length:	65047
Score:	45.00	Matches:	8
Percent Similarity:	88.89%	Conservative:	0
Best Local Similarity:	88.89%	Mismatches:	1
Query Match:	90.00%	Indels:	0
DB:	11	Gaps:	0

US-10-014-101B-40 (1-9) x ACN44020 (1-65047)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9

Db 36758 GTTCTTGAGGTTGGGGTCCCTTTTGT 36784

RESULT 2

ID AAQ64910/C

XX AAQ64910 standard; DNA; 1509 BP.

AC AAQ64910;

XX

XX 16-OCT-2003 (revised)

DT 10-JAN-1995 (first entry)

XX

DE Thermus thermophilus carotenoid biosynthesis gene.

XX

XX Carotenoid biosynthesis; food colourant; anticancer agent; antioxidant;

KW recombinant production; ds.

XX

OS Thermus thermophilus; (strain HB27).

XX

XX Key Location/Qualifiers

FH 60..1103

FT /\*tag= a

FT /function= "carotenoid\_biosynthesis"

XX

PN JP06098774-A.

XX

XX 12-APR-1994.

PD

XX

XX 25-SEP-1992; 92JP-00256580.

PF

XX

XX 25-SEP-1992; 92JP-00256580.

PR

XX

XX (MITSUBISHI PETROCHEMICAL CO LTD.

PA

XX

XX WPI; 1994-155923/19.

DR

XX

XX P-PSDB; AAR54865.

DR

XX

XX Carotenoid biosynthesis gene from Thermus - and transformed Thermus hosts

PT to enhance carotenoid prodn. for use as food colourant, anticancer agent

FT etc.

XX

PS Claim 4; Page 8-9; 12pp; Japanese.

XX

CC DNA was isolated from T.thermophilus HB27, digested with HindIII and

CC inserted into plasmid pYK134. The plasmids were used to transform T.

CC thermophilus HB27 hosts and fragments of interest were selected by growth

CC of orange colonies on kanamycin-contg. medium. Recombinant production of

CC carotenoids (useful as antioxidants, food colourants, anticancer agents,

CC etc.) can be enhanced using the transformed hosts. (Updated on 16-OCT-

XX 2003 to standardise OS field)

SQ Sequence 1509 BP; 208 A; 560 C; 537 G; 204 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	188	Length:	1509
Score:	44.00	Matches:	7
Percent Similarity:	88.89%	Conservative:	1
Best Local Similarity:	77.78%	Mismatches:	1
Query Match:	88.00%	Indels:	0
DB:	2	Gaps:	0

US-10-014-101B-40 (1-9) x AAQ64910 (1-1509)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9

Db 61 ATTCTAGCGCGCTCGGGGACTTTTGC 35

RESULT 3

ACH75978

ID ACH75978 standard; DNA; 503 BP.

XX

AC ACH75978;

XX

XX 29-JUL-2004 (first entry)

DT

XX

XX Human genome derived single exon probe #9173.

DE

XX

XX Human; probe; ss; gene expression; single exon probe; microarray;

KW alternative splicing event; genomic alteration.

XX

OS Homo sapiens.

XX

XX US2003194704-A1.

PN

XX

XX 16-OCT-2003.

PD

XX

XX 03-APR-2002; 2002US-00029386.

PF

XX

XX 03-APR-2002; 2002US-00029386.

PR

XX

XX (PENN/) PENN S G.

PA (RANK/) RANK D R.

PA (HANZ/) HANZEL D K.

XX

PI Penn SG, Rank DR, Hanzel DK;

XX

XX WPI; 2004-119264/12.

DR

XX

XX New human genome-derived single exon nucleic acid probes useful for human

PT gene expression analysis, for identifying or characterizing alternative

PT splicing events, for assessing genomic alterations or as tools for

PT surveying tissues.

XX

PS Claim 15; SEQ ID NO 9173; 80pp; English.

XX

CC The invention relates to a nucleic acid probe for measuring human gene

CC expression, comprising any of the 27,400 fully defined nucleotide

CC sequences in the specification, or their complements or fragments, and

CC encoding at least 8 amino acids of any of the 688 amino acid sequences

CC fully defined in the specification. The probe is a single exon probe that

CC hybridises under high stringency conditions to a nucleic acid molecule

CC expressed in human cells or tissues. Also included are a spatially-

CC addressable set of single exon nucleic acid probes for measuring human

CC gene expression (comprising a plurality of single exon nucleic acid

CC probes cited above, where each of the plurality of probes is separately  
 CC and addressably isolatable or amplifiable to a peptide cited above,  
 CC exon microarray for measuring human gene expression, a method of  
 CC measuring human gene expression, a vector comprising the single exon  
 CC probe cited above, an ORF-encoded peptide comprising at least 8  
 CC contiguous amino acids of any of the above-mentioned amino acid  
 CC sequences (optionally with conservative amino acid substitutions), an  
 CC isolated antibody that binds specifically to a peptide cited above,  
 CC methods of selling and/or licensing single exon probes or microarrays to  
 CC a customer desiring to measure gene expression, a method of providing  
 CC human gene expression data by subscription, and a computer-readable  
 CC storage medium which contains a database having a plurality of records  
 CC (each record including data on the expression of a single exon probe  
 CC cited above. The probe, methods and apparatus are useful in gene  
 CC expression analysis. The probes may be used as tools for surveying  
 CC tissues to detect the presence of expressed messages that contain their  
 CC specific exon, or in constructing genome-derived single exon microarrays.  
 CC In addition, the probes are used in identifying and characterising  
 CC alternative splicing events, in detecting and characterising gross  
 CC alterations in the genomic locus that includes their exon, in assessing  
 CC smaller genomic alterations, in priming the synthesis of nucleic acids,  
 CC or in expressing the ORF-encoded peptide. The present sequence is a human  
 CC single exon probe of the invention. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html?DocID=20030194704  
 XX  
 SQ Sequence 503 BP; 101 A; 124 C; 121 G; 157 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 86.4 Length: 503  
 Score: 43.00 Matches: 8  
 Percent Similarity: 88.89% Conservative: 0  
 Best Local Similarity: 88.89% Mismatches: 1  
 Query Match: 86.00% Indels: 0  
 DB: 12 Gaps: 0

US-10-014-101B-40 (1-9) x ACH75978 (1-503)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
 |||||  
 Db 4 GTTCTGGGTGGCTGGCCCTGTTCTGC 30

RESULT 4  
 ACH89659  
 ID ACH89659 standard; DNA; 789 BP.  
 AC ACH89659;  
 XX  
 XX 29-JUL-2004 (first entry)  
 DT  
 XX Human genome derived single exon probe #22854.  
 DE  
 XX Human; probe; ss; gene expression; single exon probe; microarray;  
 KW alternative splicing event; genomic alteration.  
 XX  
 OS Homo sapiens.  
 XX  
 XX US2003194704-A1.  
 PN  
 XX 16-OCT-2003.  
 PD  
 XX 03-APR-2002; 2002US-00029386.  
 PF  
 XX 03-APR-2002; 2002US-00029386.  
 PR  
 XX (PENN/) PENN S G.  
 PA (RANK/) RANK D R.  
 PA (HANZ/) HANZEL D K.  
 XX  
 XX Penn SG, Rank DR, Hanzel DK;  
 PI  
 XX WPI; 2004-119264/12.  
 DR

XX New human genome-derived single exon nucleic acid probes useful for human  
 PT gene expression analysis, for identifying or characterizing alternative  
 PT splicing events, for assessing genomic alterations or as tools for  
 PT surveying tissues.  
 XX  
 PS Claim 1; SEQ ID NO 22854; 80pp; English.  
 XX

XX The invention relates to a nucleic acid probe for measuring human gene  
 CC expression, comprising any of the 27,400 fully defined nucleotide  
 CC sequences in the specification, or their complements or fragments, and  
 CC encoding at least 8 amino acids of any of the 6888 amino acid sequences  
 CC fully defined in the specification. The probe is a single exon probe that  
 CC hybridises under high stringency conditions to a nucleic acid molecule  
 CC expressed in human cells or tissues. Also included are a spatially-  
 CC addressable set of single exon nucleic acid probes for measuring human  
 CC gene expression (comprising a plurality of single exon nucleic acid  
 CC probes cited above, where each of the plurality of probes is separately  
 CC and addressably isolatable or amplifiable from the plurality), a single  
 CC exon microarray for measuring human gene expression, a method of  
 CC measuring human gene expression, a vector comprising the single exon  
 CC probe cited above, an ORF-encoded peptide comprising at least 8  
 CC contiguous amino acids of any of the above-mentioned amino acid  
 CC sequences (optionally with conservative amino acid substitutions), an  
 CC isolated antibody that binds specifically to a peptide cited above,  
 CC methods of selling and/or licensing single exon probes or microarrays to  
 CC a customer desiring to measure gene expression, a method of providing  
 CC human gene expression data by subscription, and a computer-readable  
 CC storage medium which contains a database having a plurality of records  
 CC (each record including data on the expression of a single exon probe  
 CC cited above. The probe, methods and apparatus are useful in gene  
 CC expression analysis. The probes may be used as tools for surveying  
 CC tissues to detect the presence of expressed messages that contain their  
 CC specific exon, or in constructing genome-derived single exon microarrays.  
 CC In addition, the probes are used in identifying and characterising  
 CC alternative splicing events, in detecting and characterising gross  
 CC alterations in the genomic locus that includes their exon, in assessing  
 CC smaller genomic alterations, in priming the synthesis of nucleic acids,  
 CC or in expressing the ORF-encoded peptide. The present sequence is a human  
 CC single exon probe of the invention. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html?DocID=20030194704  
 XX

SQ Sequence 789 BP; 187 A; 176 C; 172 G; 254 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 141 Length: 789  
 Score: 43.00 Matches: 8  
 Percent Similarity: 88.89% Conservative: 0  
 Best Local Similarity: 88.89% Mismatches: 1  
 Query Match: 86.00% Indels: 0  
 DB: 12 Gaps: 0

US-10-014-101B-40 (1-9) x ACH89659 (1-789)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
 |||||  
 Db 468 GTTCTGGGTGGCTGGCCCTGTTCTGC 494

RESULT 5  
 ABK17335/c  
 ID ABK17335 standard; cDNA; 5849 BP.  
 XX  
 AC ABK17335;  
 XX

XX 26-MAR-2002 (first entry)  
 DT  
 XX Human RB-interacting zinc finger (RIZ) cDNA.  
 DE  
 XX Rat; human; retinoblastoma protein-interacting zinc finger protein; RIZ;  
 KW ss; PRDI-BF1; human positive regulatory domain 1-binding factor 1; RIZ-1;  
 KW human ecotropic viral integration site-1 myeloid transforming gene; Rb;  
 DR

KW Caenorhabditis elegans; egl-43; retinoblastoma protein; cardiac muscle;  
 KW cell proliferation; cancer; cell proliferation; neural cell; paralysis;  
 KW neurodegenerative disease; Parkinson's disease; Huntington's disease;  
 KW Alzheimer's disease; motor neuron disorder; PCR primer; mouse; EIA;  
 KW mutagenesis primer.

XX Homo sapiens.

OS US6323335-B1.

XX 27-NOV-2001.

XX 01-JUN-2000; 2000US-00586472.

XX 18-AUG-1994; 94US-00292683.

XX 06-MAR-1995; 95US-00399411.

XX 18-AUG-1995; 95US-00516859.

XX 17-MAR-2000; 2000US-00528706.

XX (HUAN/) HUANG S.

XX Huang S;

XX WPI; 2002-096600/13.

XX P-PSDB; AAU80804.

XX Novel nucleic acid molecule comprises the code for a conserved domain of

XX retinoblastoma protein-interacting zinc finger protein, positive

XX regulatory domain I-binding factor involved in regulating gene

XX transcription.

XX Example 1; Fig 9A; 93pp; English.

XX The invention relates to a nucleic acid molecule encoding a PR domain

XX peptide present in human and rat retinoblastoma protein-interacting zinc

XX finger protein (RIZ). PRI-BP1 (human positive regulatory domain I-

XX binding factor 1), EVI-1 (human ecotropic viral integration site-1

XX myeloid transforming gene product) or Caenorhabditis elegans egl-43 gene

XX product. The DNA of the invention is involved in regulating transcription

XX of a target gene. RIZ acts as a cell-differentiation factor and modulates

XX a function of a cell by binding to retinoblastoma (Rb) protein, which is

XX involved in regulating cell proliferation in various human diseases or

XX conditions, e.g. cancer. RIZ regulates the growth of normal adult cardiac

XX muscle cells and prevents proliferation of surviving cells following

XX cardiac muscle cell death. RIZ is expressed in neural cells. This is

XX useful for healing after injury of neural tissue and for treating

XX neurodegenerative diseases such as Parkinson's, Huntington's or

XX Alzheimer's disease and paralysis, or motor neuron disorders through

XX induced or decreased proliferation of neural cells. The PR domain or RIZ

XX active fragment containing a PR domain is useful as a probe to identify

XX transcription factors or oncogenic proteins in a cell that bind the PR

XX domain. Sequences ABK17334-ABK17353 represent DNA molecules and primers

XX of the invention

XX SQ Sequence 5849 BP; 1693 A; 1468 C; 1347 G; 1341 T; 0 U; 0 Other;

XX Alignment Scores:

XX Pred. No.: 1.27e+03 Length: 5849

XX Score: 43.00 Matches: 8

XX Percent Similarity: 88.89% Conservative: 0

XX Best Local Similarity: 88.89% Mismatches: 1

XX Query Match: 86.00% Indels: 0

XX DB: 6 Gaps: 0

XX US-10-014-101b-40 (1-9) x ABK17335 (1-5849)

XX QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9

XX Db 1727 GTTCTGGTGGCTGGCTGTTCTGC 1701

XX RESULT 6

XX AAT18020/c

XX ID AAT18020 standard; cDNA; 5868 BP.

XX AAT18020;  
 AC 19-MAY-1996 (first entry)  
 DT Human RIZ allele D283 cDNA.  
 XX Retinoblastoma protein-interacting zinc finger; RIZ; cell proliferation;  
 KW tumour; cancer; neuroblastoma; melanoma; diagnosis; gene therapy; ss.  
 KW Homo sapiens.  
 OS Key Location/Qualifiers  
 FH CDS 121..5280  
 FT /\*tag= a  
 FT repeat\_region 952..981  
 FT /\*tag= b  
 FT /note= "triplet repeat region is (GAA)5(GAT)5 in the D283  
 FT allele and (GAA)6(GAT)4 in the E283 allele"  
 FT 969  
 FT /\*tag= c  
 FT /note= "base 969 is t in the D283 allele and a in the  
 FT E283 allele"  
 XX WO9606168-A2.  
 XX 29-FEB-1996.  
 XX 18-AUG-1995; 95WO-US010574.  
 XX 18-AUG-1994; 94US-00292683.  
 XX 06-MAR-1995; 95US-00399411.  
 XX (LJOL-) LA JOLLA CANCER RES FOUND.  
 XX Huang S;  
 XX WPI; 1996-151371/15.  
 XX P-PSDB; AAR92100.  
 XX Nucleic acid encoding mammalian retinoblastoma protein-interacting zinc  
 PT finger - used to identify cell proliferation modulating agents for  
 PT treatment of tumours, esp. neuroblastoma and melanoma, also used for  
 PT cancer diagnosis.  
 XX Claim 4; Fig 9A; 142pp; English.  
 XX A complete human retinoblastoma (Rb) protein-interacting zinc finger  
 CC (RIZ) cDNA sequence (AAT18020) was cloned from human foetal brain cDNA  
 CC using rat RIZ cDNA (see AAT18023) coding region as probe. The human RIZ  
 CC genomic DNA has 8 exons and is located on chromosome 1p36. The RIZ D283  
 CC allele occurs 2 times more frequently in the human population than  
 CC another identified RIZ allele, E283. Clones (AAT18021-22) encoding 2  
 CC alternative 5'-termini of human RIZ were also obtd. RIZ-encoding  
 CC sequences are used for production of RIZ or its active fragments by  
 CC recombinant DNA methods, in the diagnosis of disorders associated with  
 CC abnormal levels of RIZ expression, or may be incorporated into e.g. viral  
 CC vectors for use in gene therapy

SQ Sequence 5868 BP; 1697 A; 1475 C; 1354 G; 1342 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 1.27e+03 Length: 5868  
 Score: 43.00 Matches: 8  
 Percent Similarity: 88.89% Conservative: 0  
 Best Local Similarity: 88.89% Mismatches: 1  
 Query Match: 86.00% Indels: 0  
 DB: 2 Gaps: 0

US-10-014-101b-40 (1-9) x AAT18020 (1-5868)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9

Db 1727 GTTCTGGTGGCTGGCTGTTCTGC 1701

RESULT 6

AAT18020/c

ID AAT18020 standard; cDNA; 5868 BP.

Db 1728 GTTCTGGTGGCTGGGCGCTGTTCTGC 1702

RESULT 7  
AAA60123/C  
ID AAA60123 standard; DNA; 5868 BP.

XX AC AAA60123;  
XX 17-JAN-2001 (first entry)  
XX Human RIZ allele E283 coding sequence.

XX Human; Rb-interacting zinc finger; RIZ; retinoblastoma; heart disease;  
KW cell proliferation; cell differentiation; tissue repair;  
KW transcription regulator; breast cancer; gene therapy; melanoma;  
KW neuroblastoma; leukaemia; Parkinson's disease; Huntingdon's disease;  
KW Alzheimer's disease; paralysis; motor neurone disorder; chromosome 1p36;  
ds.

XX Homo sapiens.

XX Key Location/Qualifiers  
FH CDS 121..5280  
FT /\*tag= a  
FT /product= "Human RIZ protein"  
FT repeat\_region 952..981  
FT /\*tag= b  
FT allele /rpt\_type= TANDEM  
FT replace(969,T)  
FT /\*tag= c

XX US6069231-A.  
XX 30-MAY-2000.  
XX 18-AUG-1995; 95US-00516859.  
XX 18-AUG-1994; 94US-00292683.  
XX 06-MAR-1995; 95US-00399411.  
XX (LJOL-) LA JOLLA CANCER RES FOUND.  
XX Huang S;  
XX WPI; 2000-410879/35.  
XX P-PSDB; AAB12113.

XX New PR domain peptides comprising amino acid sequences from, for example  
PT retinoblastoma-interacting zinc finger, or egl-43 proteins, for  
PT regulating gene transcription and controlling cell proliferation and  
PT differentiation.

XX Disclosure; Page; 91pp; English.

XX The present sequence is the human retinoblastoma (Rb)-interacting zinc  
CC finger (RIZ) protein allele E283 coding sequence. RIZ is a nuclear  
CC phosphoprotein that acts as a cell differentiation factor. RIZ can  
CC modulate cell growth by binding to Rb protein, which is involved in  
CC regulating cell proliferation. In addition, RIZ can act to regulate  
CC transcription. RIZ functions to maintain cells in the G1 phase of the  
CC cell cycle, by interacting with Rb through the cr2 domain of RIZ. RIZ  
CC protein is a PR domain protein and is present primarily in the cell  
CC nucleus. RIZ gene mutations may be implicated in various cancers such as  
CC melanoma, neuroblastoma, leukaemia and breast cancer, and so the RIZ gene  
CC may be used in gene therapy for these disorders. Since RIZ protein is  
CC implicated in cell cycle arrest, inhibition of RIZ activity may be useful  
CC in neurodegenerative disorder therapy e.g. for Parkinson's, Huntingdon's  
CC or Alzheimer's disease, paralysis or motor neurone disorders, or cardiac  
CC disorder therapy e.g. heart disease where the ability to induce neural/  
CC cardiac tissue proliferation would be useful. The human RIZ gene is  
CC located on chromosome 1p36. The RIZ E283 allele is thought to occur two  
CC times more frequently in the human population than the RIZ E283 allele  
CC (AAA60104 and AAB12029)

XX SQ Sequence 5868 BP; 1698 A; 1474 C; 1353 G; 1343 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 1.27e+03 Length: 5868  
Score: 43.00 Matches: 8  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 1  
Query Match: 86.00% Indels: 0  
DB: Gaps: 0

US-10-014-101B-40 (1-9) x AAA60123 (1-5868)

Qy 1 ValLeuGlyClyLeuGlyGlnPheCys 9  
Db 1728 GTTCTGGTGGCTGGGCGCTGTTCTGC 1702

RESULT 8  
AAA60104/C  
ID AAA60104 standard; DNA; 5868 BP.

XX AC AAA60104;  
XX 17-JAN-2001 (first entry)  
XX Human RIZ allele D283 coding sequence.

XX Human; Rb-interacting zinc finger; RIZ; retinoblastoma; heart disease;  
KW cell proliferation; cell differentiation; tissue repair;  
KW transcription regulator; breast cancer; gene therapy; melanoma;  
KW neuroblastoma; leukaemia; Parkinson's disease; Huntingdon's disease;  
KW Alzheimer's disease; paralysis; motor neurone disorder; chromosome 1p36;  
ds.

XX Homo sapiens.

XX Key Location/Qualifiers  
FH CDS 121..5280  
FT /\*tag= a  
FT /product= "Human RIZ protein"  
FT repeat\_region 952..981  
FT /\*tag= b  
FT allele /rpt\_type= TANDEM  
FT replace(969,A)  
FT /\*tag= c

XX US6069231-A.  
XX 30-MAY-2000.  
XX 18-AUG-1995; 95US-00516859.  
XX 18-AUG-1994; 94US-00292683.  
XX 06-MAR-1995; 95US-00399411.  
XX (LJOL-) LA JOLLA CANCER RES FOUND.  
XX Huang S;  
XX WPI; 2000-410879/35.  
XX P-PSDB; AAB12029.

XX New PR domain peptides comprising amino acid sequences from, for example  
PT retinoblastoma-interacting zinc finger, or egl-43 proteins, for  
PT regulating gene transcription and controlling cell proliferation and  
PT differentiation.

XX Example 1; Fig 9; 91pp; English.

XX The present sequence is the human retinoblastoma (Rb)-interacting zinc  
CC finger (RIZ) protein allele D283 coding sequence. RIZ is a nuclear  
CC phosphoprotein that acts as a cell differentiation factor. RIZ can  
CC modulate cell growth by binding to Rb protein, which is involved in

CC regulating cell proliferation. In addition, RIZ can act to regulate  
 CC transcription. RIZ functions to maintain cells in the G1 phase of the  
 CC cell cycle, by interacting with Rb through the cr2 domain of RIZ. RIZ  
 CC protein is a PR domain protein and is present primarily in the cell  
 CC nucleus. RIZ gene mutations may be implicated in various cancers such as  
 CC melanoma, neuroblastoma, leukaemia and breast cancer, and so the RIZ gene  
 CC may be used in gene therapy for these disorders. Since RIZ protein is  
 CC implicated in cell cycle arrest, inhibition of RIZ activity may be useful  
 CC in neurodegenerative disorder therapy e.g. for Parkinson's, Huntington's  
 CC or Alzheimer's disease, paralysis or motor neurone disorders, or cardiac  
 CC disorder therapy e.g. heart disease where the ability to induce neural/  
 CC cardiac tissue proliferation would be useful. The human RIZ gene is  
 CC located on chromosome 1p36. The RIZ D283 allele is thought to occur two  
 CC times more frequently in the human population than the RIZ E283 allele  
 CC (AAA60123 and AAB12113)

SQ Sequence 5868 BP; 1697 A; 1474 C; 1353 G; 1344 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 1.27e+03 Length: 5868  
 Score: 43.00 Matches: 8  
 Percent Similarity: 88.89% Conservative: 0  
 Best Local Similarity: 88.89% Mismatches: 1  
 Query Match: 86.00% Indels: 0  
 DB: 3 Gaps: 0

US-10-014-101B-40 (1-9) x AAA60104 (1-5868)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
 Db 1728 GTTCTGGGTGGCGCTGGCTGTTCTGC 1702

RESULT 9

AAA60124/C  
 ID AAA60124 standard; DNA; 5868 BP.

AC AAA60124;

XX 17-JAN-2001 (first entry)

XX Human mutant RIZ allele D283 coding sequence.

XX Human; Rb-interacting zinc finger; RIZ; retinoblastoma; heart disease;  
 XX cell proliferation; cell differentiation; tissue repair;  
 XX transcription regulator; breast cancer; gene therapy; melanoma;  
 XX neuroblastoma; leukaemia; Parkinson's disease; Huntington's disease;  
 XX Alzheimer's disease; paralysis; motor neurone disorder; chromosome 1p36;  
 XX mutation; ds.

XX Homo sapiens.  
 OS Synthetic.

XX Key Location/Qualifiers  
 XX CDS 121..5280  
 XX /\*tag= a  
 FT mutation /product= "Human RIZ protein"  
 FT replace(437,G)  
 FT /\*tag= b  
 FT repeat\_region 952..981  
 FT /\*tag= c  
 FT allele /rpt\_type= TANDEM  
 FT replace(969,A)  
 FT /\*tag= d

XX US6069231-A.

XX 30-MAY-2000.

XX 18-AUG-1995; 95US-00516859.

XX 18-AUG-1994; 94US-00292683.

XX 06-MAR-1995; 95US-00399411.

PA (LJOL-) LA JOLLA CANCER RES FOUND.

XX Huang S;

XX WPI; 2000-410879/35.

DR P-PSDB; AAB12114.

XX New PR domain peptides comprising amino acid sequences from, for example  
 PT retinoblastoma-interacting zinc finger, or egl-43 proteins, for  
 PT regulating gene transcription and controlling cell proliferation and  
 PT differentiation.

XX Disclosure; Page; 91pp; English.

XX The present sequence is a mutant human retinoblastoma (Rb)-interacting  
 CC zinc finger (RIZ) protein allele D283 coding sequence. RIZ is a nuclear  
 CC phosphoprotein that acts as a cell differentiation factor. RIZ can  
 CC modulate cell growth by binding to Rb protein, which is involved in  
 CC regulating cell proliferation. In addition, RIZ can act to regulate  
 CC transcription. RIZ functions to maintain cells in the G1 phase of the  
 CC cell cycle, by interacting with Rb through the cr2 domain of RIZ. RIZ  
 CC protein is a PR domain protein and is present primarily in the cell  
 CC nucleus. RIZ gene mutations may be implicated in various cancers such as  
 CC melanoma, neuroblastoma, leukaemia and breast cancer, and so the RIZ gene  
 CC may be used in gene therapy for these disorders. Since RIZ protein is  
 CC implicated in cell cycle arrest, inhibition of RIZ activity may be useful  
 CC in neurodegenerative disorder therapy e.g. for Parkinson's, Huntington's  
 CC or Alzheimer's disease, paralysis or motor neurone disorders, or cardiac  
 CC disorder therapy e.g. heart disease where the ability to induce neural/  
 CC cardiac tissue proliferation would be useful. The human RIZ gene is  
 CC located on chromosome 1p36. The RIZ D283 allele is thought to occur two  
 CC times more frequently in the human population than the RIZ E283 allele  
 CC (AAA60123 and AAB12113). Note: the present sequence is not shown in the  
 CC specification but is derived from the human RIZ allele D283 sequence  
 CC given in Figure 9 (see AAA60104)

SQ Sequence 5868 BP; 1698 A; 1474 C; 1352 G; 1344 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.27e+03 Length: 5868  
 Score: 43.00 Matches: 8  
 Percent Similarity: 88.89% Conservative: 0  
 Best Local Similarity: 88.89% Mismatches: 1  
 Query Match: 86.00% Indels: 0  
 DB: 3 Gaps: 0

US-10-014-101B-40 (1-9) x AAA60124 (1-5868)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9

Db 1728 GTTCTGGGTGGCGCTGGCTGTTCTGC 1702

RESULT 10

AAS18781/C

ID AAS18781 standard; DNA; 5868 BP.

XX AAS18781;

XX 26-MAR-2002 (first entry)

XX DNA sequence encoding human RIZ1 (hRIZ1).

XX Suppression of tumour growth; MSI positive tumour; hRIZ1; human;  
 XX retinoblastoma protein-interacting zinc finger; cancer; ds.

XX Homo sapiens.

XX Key Location/Qualifiers  
 XX CDS 121..5280  
 XX /\*tag= a  
 XX /\*product= "hRIZ1"

XX CN1313130-A.

```

XX 19-SEP-2001.
PD XX
XX 23-MAR-2001; 2001CN-00109921.
PF XX
XX 23-MAR-2001; 2001CN-00109921.
PR XX
XX 23-MAR-2001; 2001CN-00109921.
PA (HUAN/) HUANG S.
XX PI
XX PI Huang S;
XX PI
XX WPI: 2002-042166/06.
DR P-PSDB; AAU10788.
DR P-PSDB; AAU10788.
XX PT
XX Application of RIZ gene in detecting and treating tumor showing MSI
XX positive.
XX Disclosure; Page 29-36 (Disclosure); 54pp; Chinese.
XX The present invention relates to a method for suppressing growth of
XX tumours showing MSI positive. The DNA sequence of coded functional RIZ1
XX (retinoblastoma protein-interacting zinc finger) polypeptide is
XX transferred to the MSI positive tumour and the RIZ1 protein which can
XX suppress growth of the tumour is expressed in the tumour. A method for
XX determining the MSI state of a tumour is also described. The method
XX features that the number of adenines in polyadenine in the RIZ1 gene is
XX detected to determine if the tumour is MSI positive. When the number of
XX adenines is abnormal, the tumour is MSI positive. The invention is useful
XX in the treatment of cancers. The present sequence encodes for human RIZ1
XX (hRIZ1)
XX SQ Sequence 5868 BP; 1697 A; 1474 C; 1353 G; 1344 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.27e+03 Length: 5868
Score: 43.00 Matches: 8
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 1
Query Match: 86.00% Indels: 0
DB: 6 Gaps: 0

US-10-014-101B-40 (1-9) x AAS18781 (1-5868)
OY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9
DB 1728 GTCTGGTGGCTGGCCTGTTCTGC 1702

RESULT 11
ACA62689/c
ID ACA62689 standard; DNA; 5868 BP.
XX AC
XX ACA62689;
XX AC
XX 20-AUG-2003 (first entry)
XX DE Human retinoblastoma protein interacting zinc finger gene, RIZ1, DNA.
XX KW Human; gene; RIZ1; microsatellite instability; MSI; tumour; apoptosis;
XX KW retinoblastoma protein interacting zinc finger gene; colorectal tumour;
XX KW endometrial tumour; hereditary nonpolyposis colon carcinoma; ds;
XX KW gastric tumour.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX CDS 121..5280
XX FT /*tag= a
XX FT /product= "RIZ1"
XX FT replace(969,A)
XX FT /*tag= b
XX PN US2003032606-A1.
XX XX

13-FEB-2003.
PD XX
XX 17-DEC-2001; 2001US-00024450.
PF XX
XX 19-DEC-2000; 2000US-0256582P.
PR XX
XX (HUAN/) HUANG S.
PA (CHAD/) CHADWICK R B.
XX PI
XX PI Huang S, Chadwick RB;
XX PI
XX WPI: 2003-492075/46.
DR P-PSDB; ABU62110, ABU62116.
DR P-PSDB; ABU62110, ABU62116.
XX PT
XX Inhibiting growth of microsatellite instability-positive tumor, by
XX introducing a nucleic acid molecule encoding a retinoblastoma protein-
XX interacting zinc finger gene-1 polypeptide into the tumor.
XX Disclosure; Page 20-26; 41pp; English.
XX The invention relates to a method of inhibiting growth of a
XX microsatellite instability (MSI)-positive tumour, which involves
XX introducing into the tumour a nucleic acid molecule encoding a
XX retinoblastoma protein-interacting zinc finger gene (RIZ)-1 polypeptide
XX and expressing the polypeptide in the tumour in an effective amount to
XX inhibit growth of the tumour. The method is useful for inhibiting growth
XX of a microsatellite instability (MSI)-positive tumour. The tumour
XX contains cells having an abnormal number of adenosine nucleotides in a
XX RIZ poly(A) tract. The MSI-positive tumour is colorectal tumour, gastric
XX tumour, endometrial tumour or hereditary nonpolyposis colon carcinoma.
XX Also disclosed is a method for determining MSI status of the tumour. Both
XX methods are useful for detecting and treating MSI(+) tumours and for
XX inducing apoptotic cell killing both in vitro and in vivo. The present
XX sequence represents the human retinoblastoma protein interacting zinc
XX finger gene, RIZ1, DNA
XX SQ Sequence 5868 BP; 1697 A; 1474 C; 1353 G; 1344 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.27e+03 Length: 5868
Score: 43.00 Matches: 8
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 1
Query Match: 86.00% Indels: 0
DB: 9 Gaps: 0

US-10-014-101B-40 (1-9) x ACA62689 (1-5868)
OY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9
DB 1728 GTCTGGTGGCTGGCCTGTTCTGC 1702

RESULT 12
ADE84612/c
ID ADE84612 standard; cDNA; 5868 BP.
XX AC
XX ADE84612;
XX AC
XX 29-JAN-2004 (first entry)
XX DT Human retinoblastoma protein-interacting zinc finger (RIZ) protein cDNA.
XX DE cytosolic; tumor; gene therapy;
XX KW retinoblastoma protein-interacting zinc protein; RIZ; neuroblastoma;
XX KW melanoma; cardiac cell growth inducer; neuronal cell growth inducer;
XX KW human; gene; ss.
XX OS Homo sapiens.
XX PN US6468985-B1.
XX PN US6468985-B1.
XX PD 22-OCT-2002.
XX XX

```

PF 17-MAR-2000; 2000US-00528706.  
 XX 18-AUG-1994; 94US-00292683.  
 PR 06-MAR-1995; 95US-00399411.  
 PR 18-AUG-1995; 95US-00516859.  
 XX (BURN-) BURNHAM INST.  
 PA Huang S;  
 XX WPI; 2003-147106/14.  
 DR P-PSDB; ADE84613.  
 XX  
 PT Reducing the growth of tumor cells having mammalian retinoblastoma protein  
 PT -interacting zinc finger protein in a subject, comprises administering  
 PT nucleic acid molecule encoding RIZ at or adjacent to the site of tumor.  
 XX  
 XX Example 1; SEQ ID NO 3; 49pp; English.  
 XX  
 CC The invention describes a method of reducing growth of a tumour cell  
 CC having a mutant mammalian retinoblastoma (Rb) protein-interacting zinc  
 CC finger (RIZ) protein in a subject. The method comprises administering a  
 CC nucleic acid molecule (I) encoding RIZ at or adjacent to tumour site. (I)  
 CC is administered at the site of the tumour. The method is useful for  
 CC reducing the growth of a tumour cell e.g. neuroblastoma or melanoma cell  
 CC having mutant RIZ protein. The method is useful for inducing growth of a  
 CC cardiac cell or a neuronal cell in a subject, and for effecting normal  
 CC growth control to a tumour cell or causing differentiation of tumour  
 CC cells. This sequence encodes human retinoblastoma (Rb) protein-  
 CC interacting zinc finger (RIZ) protein.  
 XX  
 SQ Sequence 5868 BP; 1697 A; 1475 C; 1354 G; 1342 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 1-27e+03 Length: 5868  
 Score: 43.00 Matches: 8  
 Percent Similarity: 88.89% Conservative: 0  
 Best Local Similarity: 88.89% Mismatches: 1  
 Query Match: 86.00% Indels: 0  
 DB: 10 Gaps: 0  
 US-10-014-101B-40 (1-9) x ADE84612 (1-5868)  
 QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
 Db 1728 GTTCTGGTGGCTGGCTGGCTGTTCTGC 1702  
 RESULT 13  
 ABL33822  
 ID ABL33822 standard; DNA; 6113 BP.  
 XX  
 AC ABL33822;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Human immune system associated gene SEQ ID NO: 1795.  
 XX  
 KW Human; immune system disease; cytosine methylation; antiasthmatic;  
 KW antiarteriosclerotic; anti-anaemic; cytosine methylation; antiasthmatic;  
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
 KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;  
 KW antineuritic; cancer; eye disease; arteriosclerosis; anaemia;  
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;  
 ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200200928-A2.  
 XX  
 PD 03-JAN-2002.  
 XX  
 PF 02-JUL-2001; 2001WO-EP007537.  
 XX  
 XX 30-JUN-2000; 2000DE-01032529.  
 PR 01-SEP-2000; 2000DE-01043826.  
 XX (EPIG-) EPIGENOMICS AG.

XX 30-JUN-2000; 2000DE-01032529.  
 PR 01-SEP-2000; 2000DE-01043826.  
 XX (EPIG-) EPIGENOMICS AG.  
 XX Olek A, Piepenbrock C, Berlin K;  
 PI WPI; 2002-130909/17.  
 DR  
 XX Nucleic acid comprising fragment of chemically modified gene, useful for  
 PT diagnosis and treatment of diseases associated with abnormal cytosine  
 PT methylation.  
 XX  
 PS Claim 1; SEQ ID NO 1795; 32pp + Sequence Listing; German.  
 XX  
 CC The present invention provides a number of human immune system associated  
 CC genes which are modified by the methylation of cytosines. The sequences  
 CC can be used in the diagnosis and treatment of immune system disorders,  
 CC including eye diseases such as retinopathy, neovascular glaucoma and  
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
 CC diseases. The present sequence is a gene of the invention  
 XX  
 SQ Sequence 6113 BP; 1854 A; 134 C; 1431 G; 2694 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 1.33e+03 Length: 6113  
 Score: 43.00 Matches: 8  
 Percent Similarity: 88.89% Conservative: 0  
 Best Local Similarity: 88.89% Mismatches: 1  
 Query Match: 86.00% Indels: 0  
 DB: 6 Gaps: 0  
 US-10-014-101B-40 (1-9) x ABL33822 (1-6113)  
 QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
 Db 5780 GTTTGGGTGGGTAGGTGTTTGT 5806  
 RESULT 14  
 ABL32681  
 ID ABL32681 standard; DNA; 6436 BP.  
 XX  
 AC ABL32681;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Human immune system associated gene SEQ ID NO: 654.  
 XX  
 KW Human; immune system disease; cytosine methylation; antiasthmatic;  
 KW antiarteriosclerotic; anti-anaemic; cytosine methylation; antiasthmatic;  
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
 KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;  
 KW antineuritic; cancer; eye disease; arteriosclerosis; anaemia;  
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;  
 ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200200928-A2.  
 XX  
 PD 03-JAN-2002.  
 XX  
 PF 02-JUL-2001; 2001WO-EP007537.  
 XX  
 XX 30-JUN-2000; 2000DE-01032529.  
 PR 01-SEP-2000; 2000DE-01043826.  
 XX  
 XX (EPIG-) EPIGENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K;  
 XX WPI; 2002-130909/17.  
 XX  
 XX Nucleic acid comprising fragment of chemically modified gene, useful for  
 PT diagnosis and treatment of diseases associated with abnormal cytosine  
 PT methylation.  
 XX  
 XX Claim 1; SEQ ID NO 654; 32pp + Sequence Listing; German.  
 XX  
 XX The present invention provides a number of human immune system associated  
 CC genes which are modified by the methylation of cytosines. The sequences  
 CC can be used in the diagnosis and treatment of immune system disorders,  
 CC including eye diseases such as retinopathy, neovascular glaucoma and  
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
 CC diseases. The present sequence is a gene of the invention  
 XX  
 XX Sequence 6436 BP; 1509 A; 261 C; 1477 G; 3189 T; 0 U; 0 Other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 1.4e+03 Length: 6436  
 Score: 43.00 Matches: 8  
 Percent Similarity: 88.89% Conservative: 0  
 Best Local Similarity: 88.89% Mismatches: 1  
 Query Match: 86.00% Indels: 0  
 DB: 6 Gaps: 0  
 US-10-014-101B-40 (1-9) x ABL32681 (1-6436)  
 QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
 DB 5764 GTTTGGTGGGTAGGTGTTTTCG 5790  
 RESULT 15  
 ADF81616/C  
 ID ADF81616 standard; DNA; 6704 BP.  
 XX  
 XX ADF81616;  
 XX  
 XX 26-FEB-2004 (first entry)  
 DT  
 XX Leukaemia-related DNA sequence #2172.  
 DE  
 XX Cytostatic; Gene therapy; leukaemia; ss.  
 XX  
 XX Unidentified.  
 OS  
 XX WO2003039443-A2.  
 PN  
 XX 15-MAY-2003.  
 PD  
 XX 04-NOV-2002; 2002WO-EP012303.  
 PF  
 XX 05-NOV-2001; 2001EP-00126244.  
 XX  
 XX 30-APR-2002; 2002EP-00009758.  
 PR  
 XX (DEK-) DEUT KREBSFORSCHUNGSZENTRUM.  
 PA (UFLU-) UNIV LUDWIG MAXIMILIANS.  
 PA (HAFE/) HAERLACH T.  
 PA (SCHO/) SCHOCH C.  
 PA (KERN/) KERN W.  
 XX  
 XX Haerlach T, Schoch C, Kern W, Kohlmann A, Schnittger S, Dugas M;  
 PI Ellis R, Brors B, Mergenthaler S;  
 XX  
 XX WPI; 2003-505037/47.  
 DR  
 XX  
 XX Determining the subtype of leukemia cells and whether a patient sample  
 PT contains leukemia cells or other cells, useful for treating leukemia,  
 PT comprises determining the expression profile of a group of markers in a  
 PT patient sample.

XX  
 PS Disclosure; SEQ ID NO 2172; 2938pp; English.  
 XX  
 CC The present invention relates to a method (M1) for determining the  
 CC subtype of leukaemia cells and whether a patient sample contains  
 CC leukaemia cells. The method comprises determining the expression profile  
 CC of a group of markers in a patient sample. The method is useful for  
 CC determining the presence of leukaemia cells, its types or subtypes, and  
 CC for the preparation of a medicament for treating leukaemia.  
 XX  
 XX Sequence 6704 BP; 1992 A; 1530 C; 1451 G; 1696 T; 0 U; 35 Other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 1.47e+03 Length: 6704  
 Score: 43.00 Matches: 8  
 Percent Similarity: 88.89% Conservative: 0  
 Best Local Similarity: 88.89% Mismatches: 1  
 Query Match: 86.00% Indels: 0  
 DB: 10 Gaps: 0  
 US-10-014-101B-40 (1-9) x ADF81616 (1-6704)  
 QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
 DB 1476 GTTCTGGGTGGCTGGGCTGTCTTCG 1450  
 RESULT 16  
 ABL32308  
 ID ABL32308 standard; DNA; 7603 BP.  
 XX  
 XX ABL32308;  
 AC  
 XX 26-MAR-2002 (first entry)  
 DT  
 XX Human immune system associated gene SEQ ID NO: 281.  
 DE  
 XX Human; immune system disease; cytosine methylation; antiasthmatic;  
 XX antiarteriosclerotic; antianaemic; cytostatic; nontropic;  
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
 KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;  
 KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;  
 KW ds.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO200200928-A2.  
 PN  
 XX 03-JAN-2002.  
 PD  
 XX 02-JUL-2001; 2001WO-EP007537.  
 PF  
 XX 30-JUN-2000; 2000DE-01032529.  
 PR  
 XX 01-SEP-2000; 2000DE-01043826.  
 XX  
 XX (EPIG-) EPIGENOMICS AG.  
 PA  
 XX Olek A, Piepenbrock C, Berlin K;  
 PI WPI; 2002-130909/17.  
 DR  
 XX Nucleic acid comprising fragment of chemically modified gene, useful for  
 PT diagnosis and treatment of diseases associated with abnormal cytosine  
 PT methylation.  
 PT  
 XX Claim 1; SEQ ID NO 281; 32pp + Sequence Listing; German.  
 PS  
 XX The present invention provides a number of human immune system associated  
 CC genes which are modified by the methylation of cytosines. The sequences  
 CC can be used in the diagnosis and treatment of immune system disorders,  
 CC including eye diseases such as retinopathy, neovascular glaucoma and  
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid

CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
CC diseases. The present sequence is a gene of the invention  
SQ Sequence 7603 BP; 1892 A; 88 C; 1729 G; 3894 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 1.69e+03 Length: 7603  
Score: 43.00 Matches: 8  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 1  
Query Match: 86.00% Indels: 0  
DB: 6 Gaps: 0  
US-10-014-101B-40 (1-9) x ABL32308 (1-7603)  
QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
Db 334 GTTTTGGGTGGTGTAGGTGTTTTTGT 360  
RESULT 17  
ABL54329  
ID ABL54329 standard; DNA; 7603 BP.  
XX AC ABL54329;  
AC ABL54329;  
DT 29-JUL-2002 (first entry)  
XX DE Chemically treated apoptosis gene #15.  
XX KW Apoptosis; HIV; Bloom syndrome; cardiopathy; neurodegenerative disorder;  
KW Herpes simplex virus; renal ischaemia; amyotrophic lateral sclerosis;  
KW cancer; ds.  
XX OS Unidentified.  
XX WO200177164-A2.  
XX PN 18-OCT-2001.  
XX PD 06-APR-2001; 2001WO-EP003969.  
XX PF 06-APR-2000; 2000DE-01019058.  
XX PR 07-APR-2000; 2000DE-01019173.  
XX PR 30-JUN-2000; 2000DE-01032529.  
XX PR 01-SEP-2000; 2000DE-01043826.  
XX PA (EPIG-) EPIGENOMICS AG.  
XX PI Olek A, Piepenbrock C, Berlin K;  
XX WPI; 2002-017444/02.  
XX CC Chemically modified sequences of genes associated with apoptosis are  
PT useful to determine methylation patterns of genomic DNA samples for  
PT diagnosis of associated diseases such as cancer.  
XX PS Claim 1; Seq ID #29; 24pp; English.  
XX CC This invention relates to chemically pre-treated DNA of genes associated  
CC with apoptosis. The nucleic acids are used to allocate patients for  
CC specific therapy for HIV infection, Bloom syndrome, cardiopathy, aging,  
CC neurodegenerative disorders, Herpes simplex virus infection, renal  
CC ischaemia, amyotrophic lateral sclerosis, solid tumours and cancers. This  
CC nucleotide sequence represents a chemically treated apoptosis gene. Even  
CC SEQ ID numbers are the complementary DNA strands to the odd SEQ ID  
CC numbers. The sequence data for this patent is not represented in the  
CC printed specification but is based on information supplied by the  
CC European patent office  
SQ Sequence 7603 BP; 1892 A; 88 C; 1729 G; 3894 T; 0 U; 0 Other;  
Alignment Scores:

Pred. No.: 1.69e+03 Length: 7603  
Score: 43.00 Matches: 8  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 1  
Query Match: 86.00% Indels: 0  
DB: 6 Gaps: 0

US-10-014-101B-40 (1-9) x ABL54329 (1-7603)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
Db 334 GTTTTGGGTGGTGTAGGTGTTTTTGT 360

RESULT 18

ABK28407

ID ABK28407 standard; DNA; 9504 BP.

XX AC ABK28407;

XX DT 23-APR-2002 (first entry)

XX DE DNA transcription associated genomic DNA #141.

XX KW DNA transcription associated gene; peptide nucleic acid; PNA-oligomer;  
KW PNA; cytosine methylation state; SNP; retroviral infection; gene; ds;  
KW single nucleotide polymorphism; adenosine deaminase deficiency; cancer;  
KW viral infection; Sezary syndrome; haematological disorder; tuberculosis;  
KW immunological disorder; Werner syndrome; developmental disorder;  
KW psoriasis; Rieger's syndrome; neurological disorder; erythropoiesis;  
KW neurodegenerative disorder; Waardenburg syndrome; Niemann-Pick disease;  
KW myelodysplastic syndrome; myocardial infarction; hypertension; arthritis;  
KW angiogenesis; congenital heart disease; HDR syndrome; gene therapy;  
KW polyglutamine disorder; solid tumour.

XX OS Unidentified.

XX WO200192565-A2.

XX PD 06-DEC-2001.

XX PF 06-APR-2001; 2001WO-EP003973.

XX PR 06-APR-2000; 2000DE-01019058.

XX PR 07-APR-2000; 2000DE-01019173.

XX PR 30-JUN-2000; 2000DE-01032529.

XX PR 01-SEP-2000; 2000DE-01043826.

XX PA (EPIG-) EPIGENOMICS AG.

XX PI Olek A, Piepenbrock C, Berlin K;

XX WPI; 2002-090046/12.

XX CC New nucleic acids or oligomers, useful for diagnosing or treating  
PT diseases associated with DNA transcription, e.g. immunological disorders,  
PT Werner syndrome, psoriasis, myocardial infarction, solid tumours or  
PT cancer.

XX PS Claim 1; SEQ ID NO 281; 32pp; English.

XX CC The invention relates to a nucleic acid, which comprises a segment of the  
CC chemically pretreated DNA of genes associated with DNA transcription from  
CC one of 346 sequences, and an oligomer, in particular an oligonucleotide  
CC or peptide nucleic acid (PNA)-oligomer that hybridises to or is identical  
CC to the chemically pretreated DNA of genes associated with DNA  
CC transcription. The set of oligomer probes are useful for detecting the  
CC cytosine methylation state and/or single nucleotide polymorphisms (SNPs)  
CC in a chemically pretreated genomic DNA. The nucleic acids are useful for  
CC diagnosing or treating diseases associated with DNA transcription  
CC (particularly with the methylation status), e.g. adenosine deaminase  
CC deficiency, viral infection, retroviral infection, Sezary syndrome,  
CC haematological disorders, immunological disorders, Werner syndrome,  
CC tuberculosis, developmental disorders, psoriasis, Rieger's syndrome,

CC neurological disorders, neurodegenerative disorders, Waardenburg  
CC syndrome, Niemann-Pick disease, myelodysplastic syndrome, myocardial  
CC infarction, hypertension, angiogenesis, erythropoiesis, congenital heart  
CC disease, HDR syndrome, arthritis, polyglutamine disorders, solid tumors  
CC or cancer. Sequences ABK28127-ABK28472 represent DNA transcription  
CC associated genomic DNA molecules of the invention. Note: The sequence  
CC data for this patent did not form part of the printed specification but  
CC was obtained in electronic format directly from the European Patent  
CC Office

XX SQ Sequence 9504 BP; 2528 A; 138 C; 1853 G; 4985 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.15e+03	Length:	9504
Score:	43.00	Matches:	8
Percent Similarity:	88.89%	Conservative:	0
Best Local Similarity:	88.89%	Mismatches:	1
Query Match:	86.00%	Indels:	0
DB:	6	Gaps:	0

US-10-014-101B-40 (1-9) x ABK28407 (1-9504)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
|||||

Db 6626 GTTTGGTGGGTAGGTGTTTTTCT 6652

RESULT 19

ACN43996

ID ACN43996 standard; DNA; 254087 BP.

XX ACN43996;

XX 18-NOV-2004 (first entry)

DT Mouse genomic sequence mCG12572.

DE Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.

KW Mus musculus.

OS WO2003073826-A2.

PN 12-SEP-2003.

PD 28-FEB-2003; 2003WO-US006235.

PF 01-MAR-2002; 2002US-00087192.

PR (SAGR-) SAGRES DISCOVERY.

PA Morris DW;

PI WPI; 2003-328604/31.

DR Recombinant nucleic acid useful for diagnosis and treatment of carcinoma

XX comprises a nucleotide sequence.

PT Claim 1; SEQ ID NO 223; Opp; English.

PS The present invention relates to novel DNA and protein sequences which

XX are associated with carcinomas. The sequences are useful for: (i) for

CC screening drug candidates; (ii) for screening of bioactive agent capable

CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of

CC a bioactive agent capable of modulating the activity of CAP; (iv) for

CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing

CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating

CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;

CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for

CC determining Carcinoma Associated (CA) gene copy number. In addition, the

CC CA genes are useful as DNA vaccines and the CAP are useful as markers of

CC carcinoma including lymphoma. The present sequence is one such CA coding

CC sequence. Note: This patent is an equivalent to basic patent

CC US2002182586A1, for which no sequence data was published

XX SQ Sequence 254087 BP; 65528 A; 53970 C; 55490 G; 73068 T; 0 U; 6031 Other;

Alignment Scores:

Pred. No.:	7.83e+04	Length:	254087
Score:	43.00	Matches:	8
Percent Similarity:	88.89%	Conservative:	0
Best Local Similarity:	88.89%	Mismatches:	1
Query Match:	86.00%	Indels:	0
DB:	11	Gaps:	0

US-10-014-101B-40 (1-9) x ACN43996 (1-254087)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
|||||

Db 241873 GTACTGGGAGGCTGGGTTGTTTTC 241899

RESULT 20

ABL11263/c

ID ABL11263 standard; cDNA; 531 BP.

XX ABL11263;

AC 26-MAR-2002 (first entry)

DT Drosophila melanogaster expressed polynucleotide SEQ ID NO 28271.

XX Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ss.

XX Drosophila melanogaster.

OS WO200171042-A2.

PN 27-SEP-2001.

PD 23-MAR-2001; 2001WO-US009231.

PF 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

PI WPI; 2001-656860/75.

DR P-PSDB; ABB67160.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell

PT interactions.

PS Claim 1; SEQ ID NO 28271; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

CC sequences (AB101840-ABL16175) and the encoded proteins (ABB57737-

CC ABB72072). The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX .SQ .Sequence 531 BP; 140 A; 142 C; 167 G; 82 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	140	Length:	531
Score:	42.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	84.00%	Indels:	0

DB: 4 Gaps: 0

US-10-014-101B-40 (1-9) x ABL11263 (1-531)

QY 3 GlyGlyLeuGlyGlnPheCys 9  
|||||

Db 188 GCGGGTGGGGCAATTCCTGC 168

RESULT 21

ADRO2181

ID ADRO2181 standard; DNA; 706 BP.

XX

AC ADRO2181;

XX

DT 23-SEP-2004 (first entry)

XX

DE A. gossypii genomic DNA PAG1548RP.

XX

XX Filamentous funghi; ds; forensic identification; gene characterisation;

KW intergenomic comparison; chromosome mapping.

XX

OS Eremothecium gossypii.

XX

PN US6239264-B1.

XX

PD 29-MAY-2001.

XX

PF 24-DEC-1997; 97US-00998416.

XX

PR 31-DEC-1996; 97CH-00000016.

XX

PA (SYNG-) SYNGENTA PARTICIPATIONS AG.

XX

PI Philippsen P, Poehlmann R, Steiner-Lange S, Mohr C, Wendland J;

PI Knechtle P, Rebischung C;

XX

DR WPI; 1998-388120/33.

XX

PT New gene for adenylate cyclase from *Ashbya gossypii* - useful for

PT generating recombinant microorganisms with alteration in gene of cAMP-

PT dependent signalling pathway for increasing production of fine chemicals.

XX

PS Example 3; SEQ ID NO 873; 632pp; English.

XX

CC The invention relates to isolated DNA molecules comprising isolated

CC genomic DNA sequences from the filamentous fungi *Ashbya gossypii*, the

CC sequences comprising ADR01309, ADR01366, ADR01367, ADR01388, ADR01428,

CC ADR01466, ADR01629, ADR01637, ADR02057, ADR02345 and ADR02369, chosen

CC from 1047 disclosed genomic sequences. Also included is a cloning vector

CC comprising a nucleotide sequence chosen from the above sequences. The

CC novel *Ashbya gossypii* genomic sequences are useful for forensic

CC identification, gene characterisation, for studying gene organisation by

CC intergenomic comparison (with *Saccharomyces cerevisiae*), identifying

CC biosynthetic genes for selectable markers, to isolate

CC promoters/terminators/centromeres, chromosome mapping, and in identifying

CC sequences unique to *Ashbya gossypii* for species identification. The

CC present sequence is an *A. gossypii* novel genomic sequence of the

CC invention.

XX

SQ Sequence 706 BP; 138 A; 185 C; 182 G; 201 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	192	Length:	706
Score:	42.00	Matches:	7
Percent Similarity:	87.50%	Conservative:	0
Best Local Similarity:	87.50%	Mismatches:	1
Query Match:	84.00%	Indels:	0
DB:	2	Gaps:	0

US-10-014-101B-40 (1-9) x ADR02181 (1-706)

QY 2 LeuGlyLeuGlyGlnPheCys 9  
|||||

Db 344 CTGGGGGGTTTGGGCAATTCCTGC 367

RESULT 22

ADQ67514/c

ID ADQ67514 standard; cDNA; 2262 BP.

XX

AC ADQ67514;

XX

DT 07-OCT-2004 (first entry)

XX

DE Novel human cDNA sequence #2487.

XX

XX ss; gene; osteopathic; neuroprotective; nootropic; antiparkinsonian;

KW cytostatic; gene therapy; diagnostic marker; morbid state; osteoporosis;

KW neurological disease; Alzheimer's disease; Parkinson's disease; dementia;

XX cancer.

OS *Homo sapiens*.

XX

PN EP1440981-A2.

XX

PD 28-JUL-2004.

XX

PF 21-JAN-2004; 2004EP-00001196.

XX

PR 21-JAN-2003; 2003JP-00102206.

PR 09-MAY-2003; 2003JP-00131392.

XX

PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX

PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;

PI Yamamoto J, Isono Y, Nagai K, Irie R;

XX

DR WPI; 2004-535376/52.

DR P-PSDB; ADQ67821.

XX

PT Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,

PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.

XX

PS Claim 1; SEQ ID NO 4675; 2449pp; English.

XX

CC The invention relates to 2495 novel polynucleotides (I) and their encoded

CC polypeptides, sequences hybridizing to these nucleotides, sequences

CC encoding partial polypeptides and sequences having 70% or 90% identity to

CC the nucleotide and protein sequences. The nucleotides and polypeptides

CC are useful as diagnostic markers or therapeutic target for the diseases

CC or morbid states. They are also useful for treating osteoporosis,

CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,

CC dementia and various cancers. This sequence corresponds to a nucleotide

CC sequence of the invention.

XX

SQ Sequence 2262 BP; 595 A; 635 C; 545 G; 487 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	685	Length:	2262
Score:	42.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	84.00%	Indels:	0
DB:	12	Gaps:	0

US-10-014-101B-40 (1-9) x ADQ67514 (1-2262)

QY 3 GlyGlyLeuGlyGlnPheCys 9  
|||||

Db 1176 GGAGGACTTGGGCAATTCCTGC 1156

RESULT 23

ABL11262/c

ID ABL11262 standard; cDNA; 2531 BP.

XX

XX ABL11262;

XX



XX DE Human gene NM\_001372, SEQ ID NO 1903.  
XX KW Human; ds; gene; pain; neuronal tissue; gene therapy;  
XX KW spinal segmental nerve injury; chronic constriction injury; CCI;  
XX KW spared nerve injury; SNI; Chung.  
XX OS Homo sapiens.  
XX FN WO2003016475-A2.  
XX PD 27-FEB-2003.  
XX PF 14-AUG-2002; 2002WO-US025765.  
XX PR 14-AUG-2001; 2001US-0312147P.  
XX PR 01-NOV-2001; 2001US-0346382P.  
XX PR 26-NOV-2001; 2001US-0333347P.  
XX PA (GEHO ) GEN HOSPITAL CORP.  
XX PA (FARB ) BAYER AG.  
XX PI Woolf C, D'urso D, Befort K, Costigan M;  
XX XX WPI; 2003-268312/26.  
XX DR GENBANK; NM\_001372.  
XX XX New composition comprising two or more isolated polypeptides, useful for  
XX PT preparing a medicament for treating pain in an animal.  
XX PS Claim 1; Page; 1017pp; English.  
XX XX The invention discloses a composition comprising two or more isolated rat  
XX CC or human polynucleotides or a polynucleotide which represents a fragment,  
XX CC derivative or allelic variation of the nucleic acid sequence. Also  
XX CC claimed are a vector comprising the novel polynucleotide, a host cell  
XX CC comprising the vector, a method for identifying a nucleotide sequence  
XX CC which is differentially regulated in an animal subjected to pain and a  
XX CC kit to perform the method, an array, a method for identifying an agent  
XX CC that increases or decreases the expression of the polynucleotide sequence  
XX CC that is differentially expressed in neuronal tissue of a first animal  
XX CC subjected to pain, a method for identifying a compound which regulates  
XX CC the expression of a polynucleotide sequence which is differentially  
XX CC expressed in an animal subjected to pain, a method for identifying a  
XX CC compound that regulates the activity of one or more of the  
XX CC polynucleotides, a method for producing a pharmaceutical composition, a  
XX CC method for identifying a compound or small molecule that regulates the  
XX CC activity in an animal of one or more of the polypeptides given in the  
XX CC specification, a method for identifying a compound useful in treating  
XX CC pain and a pharmaceutical composition comprising the one or more  
XX CC polypeptides or their antibodies. The polynucleotide or the compound that  
XX CC modulates its activity is useful for preparing a medicament for treating  
XX CC pain (e.g. spinal segmental nerve injury (SNI)), chronic constriction  
XX CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
XX CC therapy). The sequence presented is a human DNA (shown in Table 2 of the  
XX CC specification) which encodes one of the polypeptides of the invention  
XX CC which is differentially expressed during pain. Note: The sequence data  
XX CC for this patent did not form part of the printed specification, but was  
XX CC obtained in electronic form directly from WIPO at  
XX CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX SQ Sequence 12374 BP; 3200 A; 3172 C; 3279 G; 2723 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 4.39e+03 Length: 12374  
Score: 42.00 Matches: 7  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 84.00% Indels: 0  
DB: 10 Gaps: 0

US-10-014-101B-40 (1-9) x ADE56064 (1-12374)

QY 3 GlyGlyLeuGlyGlnPheCys 9  
Db 2821 GGAGGTCCTGGGTCAGTTCTGC 2841  
RESULT 26  
ADD46830  
ID ADD46830 standard; DNA; 12374 BP.  
XX AC ADD46830;  
XX AC  
XX DT 02-DEC-2004 (revised)  
XX DT 29-JAN-2004 (first entry)  
XX DE Human gene NM\_001372, SEQ ID NO 12516.  
XX KW Human; ds; gene; pain; neuronal tissue; gene therapy;  
XX KW spinal segmental nerve injury; chronic constriction injury; CCI;  
XX KW spared nerve injury; SNI; Chung.  
XX OS Homo sapiens.  
XX OS Unidentified.  
XX PN WO2003016475-A2.  
XX PD 27-FEB-2003.  
XX PF 14-AUG-2002; 2002WO-US025765.  
XX PR 14-AUG-2001; 2001US-0312147P.  
XX PR 01-NOV-2001; 2001US-0346382P.  
XX PR 26-NOV-2001; 2001US-0333347P.  
XX XX (GEHO ) GEN HOSPITAL CORP.  
XX PA (FARB ) BAYER AG.  
XX PI Woolf C, D'urso D, Befort K, Costigan M;  
XX XX WPI; 2003-268312/26.  
XX DR GENBANK; NM\_001372.  
XX XX New composition comprising two or more isolated polypeptides, useful for  
XX PT preparing a medicament for treating pain in an animal.  
XX PS Example 1; Page; 1017pp; English.  
XX XX The invention discloses a composition comprising two or more isolated rat  
XX CC or human polynucleotides or a polynucleotide which represents a fragment,  
XX CC derivative or allelic variation of the nucleic acid sequence. Also  
XX CC claimed are a vector comprising the novel polynucleotide, a host cell  
XX CC comprising the vector, a method for identifying a nucleotide sequence  
XX CC which is differentially regulated in an animal subjected to pain and a  
XX CC kit to perform the method, an array, a method for identifying an agent  
XX CC that increases or decreases the expression of the polynucleotide sequence  
XX CC that is differentially expressed in neuronal tissue of a first animal  
XX CC subjected to pain, a method for identifying a compound which regulates  
XX CC the expression of a polynucleotide sequence which is differentially  
XX CC expressed in an animal subjected to pain, a method for identifying a  
XX CC compound that regulates the activity of one or more of the  
XX CC polynucleotides, a method for producing a pharmaceutical composition, a  
XX CC method for identifying a compound or small molecule that regulates the  
XX CC activity in an animal of one or more of the polypeptides given in the  
XX CC specification, a method for identifying a compound useful in treating  
XX CC pain and a pharmaceutical composition comprising the one or more  
XX CC polypeptides or their antibodies. The polynucleotide or the compound that  
XX CC modulates its activity is useful for preparing a medicament for treating  
XX CC pain (e.g. spinal segmental nerve injury (SNI)), chronic constriction  
XX CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
XX CC therapy). The sequence presented is a human DNA (described in Table 3 of  
XX CC the specification) which encodes one of the polypeptides of the invention  
XX CC which is differentially expressed during pain. Note: The sequence data  
XX CC for this patent did not form part of the printed specification, but was  
XX CC obtained in electronic form directly from WIPO at  
XX CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 12374 BP; 3200 A; 3172 C; 3279 G; 2723 T; 0 U; 0 Other;  
 Alignment Scores: Pred. No.: 4.39e+03 Length: 12374  
 Score: 42.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 84.00% Indels: 0  
 DB: 10 Gaps: 0

US-10-014-101B-40 (1-9) x ADD46830 (1-12374)

QY 3 GlyGlyLeuGlyGlnPheCys 9  
 |||||  
 DB 2821 GGAGGTCGTGGTCAGTTCTGC 2841

RESULT 27  
 AAS46421  
 ID AAS46421 standard; DNA; 13321 BP.  
 AC AAS46421;  
 XX  
 XX 18-DEC-2001 (first entry)  
 XX  
 XX Tumour suppressor gene derived chemically modified sequence #143.  
 DE Human; tumour suppressor gene; oncogene; antitumour; cytostatic; cancer;  
 XX tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;  
 KW cytosine methylation; ds.  
 XX Homo sapiens.  
 OS  
 XX  
 XX WO200168912-A2.  
 PN  
 XX  
 XX 20-SEP-2001.  
 PD  
 XX  
 XX 15-MAR-2001; 2001WO-EP002955.  
 PF  
 XX  
 XX 15-MAR-2000; 2000DE-01013847.  
 PR  
 XX 06-APR-2000; 2000DE-01019058.  
 PR  
 XX 07-APR-2000; 2000DE-01019173.  
 PR  
 XX 30-JUN-2000; 2000DE-01032529.  
 PR  
 XX 01-SEP-2000; 2000DE-01043826.  
 PR  
 XX (EPIG-) EPIGENOMICS AG.  
 PA  
 XX  
 XX Olek A, Piepenbrock C, Berlin K;  
 PI  
 XX WPI; 2001-602752/68.  
 DR  
 XX  
 XX Fragments of chemically modified genes associated with tumor suppressor  
 PT genes and oncogenes, useful in designing primers and probes for analyzing  
 PT diseases associated with cytosine methylation state e.g. cancer.  
 XX  
 PS Claim 1; SEQ ID NO 143; 27pp; English.

The invention relates to a nucleic acid comprising a sequence of 18  
 CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with  
 CC bisulphite, of genes associated with tumour suppression and oncogenes  
 CC having a sequence taken from 536 (actually 533 since numbers 408, 458 and  
 CC 500 are missing from the sequence listing) sequences (Ss) and sequences  
 CC complementary to (Ss). The nucleic acid may be a peptide nucleic acid-  
 CC oligomer (PNA) of at least 9 nucleotides and may form part of a set of  
 CC probes for detecting the cytosine methylation state and/or single  
 CC nucleotide polymorphisms and also to be used in an array for analysing  
 CC diseases associated with CpG dinucleotides e.g. cancers and tumours. The  
 CC probes can also be used in a method for ascertaining genetic and/or  
 CC epigenetic parameters for the diagnosis and/or therapy of existing  
 CC diseases or the predisposition to specific diseases, by analysing  
 CC cytosine methylations. The parameters may be compared to another set of  
 CC genetic and/or epigenetic parameters, the differences serving as basis  
 CC for diagnosis and/or prognostic events which are disadvantageous to

CC patients. The present sequence is one of the 533 genomic sequences  
 CC derived from tumour suppressor genes and oncogenes. Note: The sequence  
 CC data for this patent did not form part of the printed specification, but  
 CC was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 13321 BP; 2515 A; 580 C; 4121 G; 6105 T; 0 U; 0 Other;  
 Alignment Scores: Pred. No.: 4.76e+03 Length: 13321  
 Score: 42.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 3  
 Best Local Similarity: 66.67% Mismatches: 0  
 Query Match: 84.00% Indels: 0  
 DB: 4 Gaps: 0

US-10-014-101B-40 (1-9) x AAS46421 (1-13321)

QY 1 ValLeuGlyLeuGlyGlnPheCys 9  
 |||||  
 DB 10259 GTTATAGTGTGTCGAGAAATTTGT 10285

RESULT 28  
 AAH82100/C  
 ID AAH82100 standard; DNA; 255 BP.  
 XX  
 XX AAH82100;  
 AC  
 XX 21-SEP-2001 (first entry)  
 DT  
 XX  
 XX Rat differential transcription-associated cDNA SEQ ID 609.  
 DE  
 XX Differential transcription; human; rat; tumour cell; cytostatic;  
 KW Ras modulator; Class II tumour suppressor gene; gene therapy; ss.  
 XX  
 XX Rattus sp.  
 OS  
 XX  
 XX WO200157058-A2.  
 PN  
 XX  
 XX 09-AUG-2001.  
 PD  
 XX  
 XX 31-JAN-2001; 2001WO-EP001003.  
 PF  
 XX  
 XX 31-JAN-2000; 2000DE-01004102.  
 PR  
 XX (META-) METAGEN GES GENOMFORSCHUNG MBH.  
 PA  
 XX  
 XX Rosenthal A, Hinzmann B, Schaefer R, Zuber J, Tchernitsa O;  
 PI Grips M, Hellriegel M, Schmitz A, Sers C;  
 PI  
 XX WPI; 2001-483415/52.  
 DR  
 XX  
 XX Nucleic acids differentially expressed between tumor and normal cells,  
 PT useful for diagnosis or therapy of tumors and for screening active  
 PT agents.

XX  
 PS Claim 6; Page 529; 579pp; German.

This invention describes a nucleic acid (I) with differential expression  
 CC between tumour and normal cells and which has cytostatic activity. (I)  
 CC work as modulators of Ras activity by inducing expression of tumour  
 CC suppressor genes. (I), and polypeptides encoded by them, are useful as  
 CC targets for diagnosis or therapy and in screening to determine the  
 CC effects of an active compound (potential pharmaceutical) on a cell line,  
 CC particularly for diagnosis and treatment of tumors, especially by  
 CC modulating expression of (I) (by gene therapy, antisense RNA or ribozyme  
 CC methods) or by modulating the amount and/or location of (I)-encoded  
 CC polypeptides (by administration of the polypeptide or its activator,  
 CC antibody (optionally as a conjugate) or inhibitor). The method allows  
 CC identification of many Class II tumour suppressor genes (i.e. genes that  
 CC are not primary targets for tumour-initiating mutations). AAH81492-  
 CC AAH82376 represent the human and rat derived nucleic acid fragments  
 CC described in the method of the invention

XX SQ Sequence 255 BP; 83 A; 45 C; 77 G; 50 T; 0 U; 0 Other;

Alignment Scores: 96.3 Length: 255  
Pred. No.: 41.00 Matches: 7  
Score: 87.50% Conservative: 0  
Percent Similarity: 87.50% Mismatches: 1  
Best Local Similarity: 82.00% Indels: 0  
Query Match: 5 Gaps: 0  
DB: 5

US-10-014-101B-40 (1-9) x AH82100 (1-255)

QY 2 LeuGlyGlyLeuGlyGlnPheCys 9  
Db 179 CTGGTGGTGGGCCATTCTGT 156

RESULT 29  
ABN23853/c  
ID ABN23853 standard; cDNA; 321 BP.  
XX AC ABN23853;  
XX AC  
XX 24-JUN-2002 (first entry)  
XX Human ORFX polynucleotide sequence SEQ ID NO:16183.  
XX Human; open reading frame; ORFX; Gene therapy; cancer; cirrhosis;  
XX hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;  
XX degenerative disorder; osteoarthritis; neurodegenerative disorder;  
XX cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;  
XX hypertension; hypothyroidism; cholesterol ester storage disease;  
XX immune deficiency; immune disorder; infectious disease;  
XX autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;  
XX myasthenia gravis; gene; ss.  
XX Homo sapiens.  
XX OS  
XX WO200192523-A2.  
XX PN  
XX 06-DEC-2001.  
XX PD  
XX 29-MAY-2001; 2001WO-US010836.  
XX PF  
XX 30-MAY-2000; 2000US-0206132P.  
XX PR  
XX 29-AUG-2000; 2000US-0228716P.  
XX XX  
XX (CURA-) CURAGEN CORP.  
XX PA  
XX Shimketa RA, Leach WD;  
XX PI  
XX WPI; 2002-106308/14.  
XX DR  
XX P-PSDB; ABP08101.  
XX DR  
XX Novel human polypeptides and polynucleotides useful for diagnosing,  
XX preventing and treating cardiovascular disease, neurodegenerative,  
XX hyperproliferative disorders and autoimmune disorders.  
XX PT  
XX Disclosure; SEQ ID NO 16183; 1037pp; English.  
XX PS  
XX The present invention describes substantially purified human proteins  
XX (referred to as open reading frame, ORFX, where X is 1-11491 (see table 1  
XX in the specification). ABN15762 to ABN27252 encode the human ORFX  
XX proteins given in ABP00010 to ABP11500. ORFX proteins are useful for  
XX treating or preventing a pathology associated with an ORFX-associated  
XX disorder in humans, and in the manufacture of a medicament for treating a  
XX syndrome associated with ORFX-associated disorder. ORFX polynucleotide  
XX sequences can be used in gene therapy. ORFX sequences can be used in the  
XX treatment of cancer, hyperproliferative disorders, cirrhosis of liver,  
XX psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,  
XX osteoarthritis, neurodegenerative disorders, disorders related to organ  
XX transplantation, cardiovascular diseases, diabetes mellitus, systemic  
XX lupus erythematosus, hypertension, hypothyroidism, cholesterol ester

CC storage disease, various immune deficiencies and disorders, infectious  
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid  
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host  
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also  
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,  
CC bone degenerative disorders, or periodontal disease, and for gut  
CC protection or regeneration and treatment of lung or liver fibrosis,  
CC reperfusion injury in various tissues and conditions resulting from  
CC systemic cytokine damage. N.B. The sequence data for this patent did not  
CC form part of the printed specification, but was obtained in electronic  
CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX SQ Sequence 321 BP; 105 A; 66 C; 74 G; 76 T; 0 U; 0 Other;

Alignment Scores: 124 Length: 321  
Pred. No.: 41.00 Matches: 6  
Score: 100.00% Conservative: 3  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 66.67% Indels: 0  
Query Match: 82.00% Gaps: 0  
DB: 6

US-10-014-101B-40 (1-9) x ABN23853 (1-321)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
Db 249 GTTATTGGTGGGTTGGAGATTCTGC 223

RESULT 30  
AAC36788  
ID AAC36788 standard; DNA; 503 BP.  
XX AC AAC36788;  
XX AC  
XX 17-OCT-2000 (first entry)  
XX DT  
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 15080.  
XX DE  
XX Arabidopsis thaliana.  
XX OS  
XX Hybridisation assay; genetic mapping; gene expression control;  
XX protein identification; signal transduction pathway; metabolic pathway;  
XX promoter; termination sequence; ss.  
XX KW  
XX Arabidopsis thaliana.  
XX EP1033405-A2.  
XX PN  
XX 06-SEP-2000.  
XX PD  
XX 25-FEB-2000; 2000EP-00301439.  
XX PF  
XX 25-FEB-1999; 99US-0121825P.  
XX PR  
XX 05-MAR-1999; 99US-0123180P.  
XX PR  
XX 09-MAR-1999; 99US-0123548P.  
XX PR  
XX 23-MAR-1999; 99US-0125788P.  
XX PR  
XX 25-MAR-1999; 99US-0126264P.  
XX PR  
XX 29-MAR-1999; 99US-0126785P.  
XX PR  
XX 01-APR-1999; 99US-0127462P.  
XX PR  
XX 08-APR-1999; 99US-0128234P.  
XX PR  
XX 16-APR-1999; 99US-0128714P.  
XX PR  
XX 19-APR-1999; 99US-0130077P.  
XX PR  
XX 21-APR-1999; 99US-0130449P.  
XX PR  
XX 23-APR-1999; 99US-0130510P.  
XX PR  
XX 28-APR-1999; 99US-0130891P.  
XX PR  
XX 30-APR-1999; 99US-0131449P.  
XX PR  
XX 30-APR-1999; 99US-0132048P.  
XX PR  
XX 04-MAY-1999; 99US-0132407P.  
XX PR  
XX 04-MAY-1999; 99US-0132484P.  
XX PR  
XX 05-MAY-1999; 99US-0132485P.  
XX PR  
XX 06-MAY-1999; 99US-0132486P.  
XX PR  
XX 07-MAY-1999; 99US-0132487P.  
XX PR  
XX 07-MAY-1999; 99US-0132863P.  
XX PR  
XX 11-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 14-MAY-1999; 99US-0134768P.  
PR 15-MAY-1999; 99US-0134941P.  
PR 15-MAY-1999; 99US-0135124P.  
PR 15-MAY-1999; 99US-0135353P.  
PR 15-MAY-1999; 99US-0135629P.  
PR 15-MAY-1999; 99US-0136021P.  
PR 15-MAY-1999; 99US-0136392P.  
PR 15-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 21-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 28-JUN-1999; 99US-0140981P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 08-JUL-1999; 99US-0142820P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143342P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 16-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 23-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.

PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 08-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-015753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.

PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161992P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:  
 Pred. No.: 202 Length: 503  
 Score: 41.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 82.00% Indels: 0  
 DB: 3 Gaps: 0

US-10-014-101B-40 (1-9) x AAC36788 (1-503)

QY 1 ValLeuGlyGlyLeuGlyGlnPhe 8  
 DB 74 GTGTAGGAGTTGGGTCAATT 97

RESULT 31  
 ACA40032/c  
 ID ACA40032 standard; DNA; 537 BP.

XX ACA40032;  
 XX

DT 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #21689.

DE Antisense; ds; prokaryotic essential gene; cell proliferation;  
 KW drug design; gene.

XX Mycoplasma pneumoniae.

XX WC200277183-A2.

PN 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

DR P-PSDB; ABU36162.

XX New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.

XX Claim 14; SEQ ID NO 27902; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway; (8)  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
 CC prokaryotic essential genes. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX  
 SQ Sequence 537 BP; 187 A; 115 C; 102 G; 133 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 217 Length: 537  
 Score: 41.00 Matches: 7  
 Percent Similarity: 88.89% Conservative: 1  
 Best Local Similarity: 77.78% Mismatches: 1  
 Query Match: 82.00% Indels: 0  
 DB: 8 Gaps: 0

US-10-014-101B-40 (1-9) x ACA40032 (1-537)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
 DB 76 GTTGTAGCGGGCTTGGTCAGTTTGC 50

RESULT 32

ADH61274

ID ADH61274 standard; DNA; 552 BP.

XX AC ADH61274;

XX DT 25-MAR-2004 (first entry)

XX DE Soybean cytochrome oxidase (CCK1) DNA #3.

XX KW Male-sterility; cytochrome oxidase; transgenic plant; seed production;  
 KW CKX1; soybean; ds; gene.

XX OS Glycine max.

XX PN US2003163847-A1.

PD 28-AUG-2003.

XX PF 20-DEC-2002; 2002US-00326194.

XX PR 20-DEC-2001; 2001US-0343129P.

XX (PHRA ) PHARMACIA CORP.

XX PI Huang S, Crossland LD, Cheikh N, Morris RO;

XX WPI; 2003-897983/82.

XX GENBANK; BM528151.

PT Producing plants characterized by reversible male-sterility, useful for

PT maintaining male sterility in plants, by transforming a plant cell with a  
 PT nucleic acid construct containing a polynucleotide encoding a cytokinin  
 PT oxidase.

XX Disclosure; SEQ ID NO 15; 33pp; English.

XX The invention relates to a method for producing a plant characterised by  
 CC reversible male-sterility which involves transforming a plant cell with a  
 CC nucleic acid construct containing a polynucleotide encoding a cytokinin  
 CC oxidase. The method is useful for producing reversible male-sterility in  
 CC transgenic plants, or for maintaining male sterility in plants. The  
 CC method reduces the expense of seed production for existing hybrid plants  
 CC such as corn, but also makes it possible to produce hybrid varieties of  
 CC traditionally non-hybrid crops. The method is also useful for introducing  
 CC economically valuable traits from plants having undesirable production  
 CC characteristics into plants having desirable characteristics. The present  
 CC sequence is soybean cytokinin oxidase (CKX1) DNA. This sequence is used  
 CC to illustrate the method of the invention.

SQ Sequence 552 BP; 167 A; 106 C; 118 G; 161 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 224 Length: 552  
 Score: 41.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 82.00% Indels: 0  
 DB: 10 Gaps: 0

US-10-014-101B-40 (1-9) x ADH61274 (1-552)

QY 1 ValLeuGlyLeuGlyGlnPhe 8  
 DB 12 GTTCTGGAGGCTTAGGGCAATTT 35

RESULT 33

ADH61273  
 ID ADH61273 standard; DNA; 566 BP.

XX AC ADH61273;

XX 25-MAR-2004 (first entry)

XX Soybean cytokinin oxidase (CKX1) DNA #2.

XX Male-sterility; cytokinin oxidase; transgenic plant; seed production;  
 XX CKX1; soybean; ds; gene.

XX Glycine max.

XX US2003163847-A1.

XX 28-AUG-2003.

XX 20-DEC-2002; 2002US-00326184.

XX 20-DEC-2001; 2001US-0343129P.

XX (PHAA ) PHARMACIA CORP.

XX Huang S, Crossland LD, Cheikh N, Morris RO;

XX WPI; 2003-897983/82.

XX GENBANK; BU084470.

XX Producing plants characterized by reversible male-sterility, useful for  
 PT maintaining male sterility in plants, by transforming a plant cell with a  
 PT nucleic acid construct containing a polynucleotide encoding a cytokinin  
 PT oxidase.

XX Disclosure; SEQ ID NO 14; 33pp; English.

XX The invention relates to a method for producing a plant characterised by

CC reversible male-sterility which involves transforming a plant cell with a  
 CC nucleic acid construct containing a polynucleotide encoding a cytokinin  
 CC oxidase. The method is useful for producing reversible male-sterility in  
 CC transgenic plants, or for maintaining male sterility in plants. The  
 CC method reduces the expense of seed production for existing hybrid plants  
 CC such as corn, but also makes it possible to produce hybrid varieties of  
 CC traditionally non-hybrid crops. The method is also useful for introducing  
 CC economically valuable traits from plants having undesirable production  
 CC characteristics into plants having desirable characteristics. The present  
 CC sequence is soybean cytokinin oxidase (CKX1) DNA. This sequence is used  
 CC to illustrate the method of the invention.

SQ Sequence 566 BP; 162 A; 115 C; 132 G; 157 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 230 Length: 566  
 Score: 41.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 82.00% Indels: 0  
 DB: 10 Gaps: 0

US-10-014-101B-40 (1-9) x ADH61273 (1-566)

QY 1 ValLeuGlyLeuGlyGlnPhe 8  
 DB 283 GTTCTGGAGGCTTAGGGCAATTT 306

RESULT 34

ADH61277

ID ADH61277 standard; DNA; 582 BP.

XX AC ADH61277;

XX 25-MAR-2004 (first entry)

XX Soybean cytokinin oxidase (CKX1) DNA #6.

XX Male-sterility; cytokinin oxidase; transgenic plant; seed production;  
 XX CKX1; soybean; ds; gene.

XX Glycine max.

XX US2003163847-A1.

XX 28-AUG-2003.

XX 20-DEC-2002; 2002US-00326184.

XX 20-DEC-2001; 2001US-0343129P.

XX (PHAA ) PHARMACIA CORP.

XX Huang S, Crossland LD, Cheikh N, Morris RO;

XX WPI; 2003-897983/82.

XX GENBANK; BM143498.

XX Producing plants characterized by reversible male-sterility, useful for  
 PT maintaining male sterility in plants, by transforming a plant cell with a  
 PT nucleic acid construct containing a polynucleotide encoding a cytokinin  
 PT oxidase.

XX Disclosure; SEQ ID NO 18; 33pp; English.

XX The invention relates to a method for producing a plant characterised by  
 CC reversible male-sterility which involves transforming a plant cell with a  
 CC nucleic acid construct containing a polynucleotide encoding a cytokinin  
 CC oxidase. The method is useful for producing reversible male-sterility in  
 CC transgenic plants, or for maintaining male sterility in plants. The  
 CC method reduces the expense of seed production for existing hybrid plants  
 CC such as corn, but also makes it possible to produce hybrid varieties of  
 CC traditionally non-hybrid crops. The method is also useful for introducing

CC economically valuable traits from plants having undesirable production  
CC characteristics into plants having desirable characteristics. The present  
CC sequence is soybean cytokinin oxidase (CKX1) DNA. This sequence is used  
CC to illustrate the method of the invention.

SQ Sequence 582 BP; 179 A; 112 C; 129 G; 162 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 237 Length: 582  
Score: 41.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 82.00% Indels: 0  
DB: 10 Gaps: 0

US-10-014-101B-40 (1-9) x ADH61277 (1-582)

QY 1 ValLeuGlyGlyLeuGlyGlnPhe 8  
DB 58 GTTCTTGGAGGCTTGGGCAATT 81

## RESULT 35

ADH61275  
ID ADH61275 standard; DNA; 635 BP.

XX AC ADH61275;

XX DT 25-MAR-2004 (first entry)

XX DE Soybean cytokinin oxidase. (CKX1) DNA #4.

XX KW Male-sterility; cytokinin oxidase; transgenic plant; seed production;  
XX CKX1; soybean; ds; gene.

XX OS Glycine max.

XX PN US2003163847-A1.

XX PD 28-AUG-2003.

XX PF 20-DEC-2002; 2002US-00326184.

XX PR 20-DEC-2001; 2001US-0343129P.

XX PA (PHAA ) PHARMACIA CORP.

XX PI Huang S, Crossland LD, Cheikh N, Morris RO;

XX DR WPI: 2003-897983/82.

XX DR GENBANK; BE330968.

XX PT Producing plants characterized by reversible male-sterility, useful for  
XX maintaining male sterility in plants, by transforming a plant cell with a  
XX nucleic acid construct containing a polynucleotide encoding a cytokinin  
XX oxidase.

XX PS Disclosure; SEQ ID NO 16; 33pp; English.

XX The invention relates to a method for producing a plant characterised by  
XX reversible male-sterility which involves transforming a plant cell with a  
XX nucleic acid construct containing a polynucleotide encoding a cytokinin  
XX oxidase. The method is useful for producing reversible male-sterility in  
XX transgenic plants, or for maintaining male sterility in plants. The  
XX method reduces the expense of seed production for existing hybrid plants  
XX such as corn, but also makes it possible to produce hybrid varieties of  
XX traditionally non-hybrid crops. The method is also useful for introducing  
XX economically valuable traits from plants having undesirable production  
XX characteristics into plants having desirable characteristics. The present  
XX sequence is soybean cytokinin oxidase (CKX1) DNA. This sequence is used  
XX to illustrate the method of the invention.

SQ Sequence 635 BP; 198 A; 122 C; 137 G; 178 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 261 Length: 635  
Score: 41.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 82.00% Indels: 0  
DB: 10 Gaps: 0

US-10-014-101B-40 (1-9) x ADH61275 (1-635)

QY 1 ValLeuGlyGlyLeuGlyGlnPhe 8  
DB 108 GTTCTTGGAGGCTTGGGCAATT 131

## RESULT 36

AAH34132

ID AAH34132 standard; cDNA; 1088 BP.

XX AC AAH34132;

XX DT 03-SEP-2001 (first entry)

XX DE Human colon cancer antigen encoding cDNA SEQ ID NO:1214.

XX KW Human; colon cancer; colon cancer antigen; diagnosis; detection;  
XX colorectal carcinoma; chromosome 15; ss.

XX OS Homo sapiens.

XX PN WO200122920-A2.

XX PD 05-APR-2001.

XX PF 28-SEP-2000; 2000WO-US026524.

XX PR 29-SEP-1999; 99US-0157137P.

XX PR 03-NOV-1999; 99US-0163280P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Ruben SM, Barash SC, Birse CE, Rosen CA;

XX DR WPI; 2001-235357/24.

XX DR P-PSDB; AAG74727.

XX PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
XX useful for preventing, diagnosing and/or treating colorectal cancers.

XX PS Claim 1; Page 2995; 9803pp; English.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon  
XX cancer-associated nucleic acid molecules (N) and proteins (P), where the  
XX proteins are collectively known as colon cancer antigens. The colon  
XX cancer antigens have cytostatic activity and can be used in gene therapy  
XX and vaccine production. N and P may be used in the prevention, diagnosis  
XX and treatment of diseases associated with inappropriate P expression. For  
XX example, N and P may be used to treat disorders associated with decreased  
XX expression by rectifying mutations or deletions in a patient's genome  
XX that affect the activity of P by expressing inactive proteins or to  
XX supplement the patient's own production of P. Additionally, N may be used  
XX to produce the colon cancer-associated Ps, by inserting the nucleic acids  
XX into a host cell and culturing the cell to express the proteins. N and P  
XX can be used in the prevention, diagnosis and treatment of colorectal  
XX carcinomas and cancers. AAH37196 to AAH37204 and AAG77789 represent  
XX sequences used in the exemplification of the present invention. N.B.  
XX Pages 666 to 682 and page 7053 of the sequence listing were missing at  
XX time of publication, meaning no sequences are present for SEQ ID NO:1027  
XX to 1052, 7921 and 7922

SQ Sequence 1088 BP; 309 A; 239 C; 261 G; 276 T; 0 U; 3 Other;

## Alignment Scores:

Pred. No.: 471 Length: 1088

Score: 41.00 Matches: 7  
Percent Similarity: 87.50% Conservative: 0  
Best Local Similarity: 87.50% Mismatches: 1  
Query Match: 82.00% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-40 (1-9) x AAH34132 (1-1088)

QY 2 LeuGlyGlyLeuGlyGlnPheCys 9

Db 59 CTCGTCGGCCTCGCTCGTTCTGC 82

RESULT 37

ABQ34200

ID ABQ34200 standard; DNA; 1234 BP.

XX AC ABQ34200;

XX DT 12-JUL-2002 (first entry)

XX DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 20791.

XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;  
KW drug; side effect; cancer; central nervous system; cardiovascular;  
KW gastrointestinal; respiratory system; single nucleotide polymorphism;  
KW SNP; cell differentiation; ds.

XX OS Homo sapiens.

XX PN WO200218632-A2.

XX PD 07-MAR-2002.

XX PF 01-SEP-2001; 2001WO-EP010074.

XX PR 01-SEP-2000; 2000DE-01043826.

XX PR 05-SEP-2000; 2000DE-01044543.

XX PA (EPIG-) EPIGENOMICS AG.

XX PI Olek A, Piepenbrock C, Berlin K, Guetig D;

XX WPI; 2002-371829/40.

XX Determining the degree of cytosine methylation in genomic DNA, useful for  
PT diagnosis and prognosis, comprises selective hybridization of amplicons  
PT from chemically treated DNA.

XX PS Claim 12; 56pp + Sequence Listing; 56pp; German.

XX This invention describes a novel method for determining the degree of  
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
CC genomic sample of DNA. The sample is treated chemically to convert  
CC cytosine (C) but not methylated C, to uracil, then part of the genomic  
CC DNA that contains the target C is amplified to form a labeled amplicon.  
CC The amplicon is hybridised to two classes, each with at least one member,  
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the  
CC degree of hybridisation to both classes is determined from the label on  
CC the amplicon. From the ratio of labels hybridised to the two classes of  
CC oligomers, the degree of methylation is calculated. The method is used:  
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs  
CC and of a wide range of diseases, e.g. cancer, disorders of the central  
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,  
CC particularly by detecting mutations or single nucleotide polymorphisms  
CC (SNP's); and (ii) for differentiation of cell or tissue types and for  
CC investigating cell differentiation. The method allows the methylation  
CC status of many C residues to be determined simultaneously. ABQ13410-  
CC ABQ54121 represent genomic DNA sequences used to illustrate the method  
CC for determining the degree of cytosine methylation described in the  
CC disclosure of the invention

XX SQ Sequence 1234 BP; 205 A; 151 C; 498 G; 380 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 540 Length: 1234  
Score: 41.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 3  
Best Local Similarity: 66.67% Mismatches: 0  
Query Match: 82.00% Indels: 0  
DB: 6 Gaps: 0

US-10-014-101B-40 (1-9) x ABQ34200 (1-1234)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9

Db 532 GTCTGCGGCGGCTTGGTCAAGTTTGC 558

RESULT 38

ABQ34201/C

ID ABQ34201 standard; DNA; 1234 BP.

XX AC ABQ34201;

XX DT 12-JUL-2002 (first entry)

XX DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 20792.

XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;  
KW drug; side effect; cancer; central nervous system; cardiovascular;  
KW gastrointestinal; respiratory system; single nucleotide polymorphism;  
KW SNP; cell differentiation; ds.

XX OS Homo sapiens.

XX PN WO200218632-A2.

XX PD 07-MAR-2002.

XX PF 01-SEP-2001; 2001WO-EP010074.

XX PR 01-SEP-2000; 2000DE-01043826.

XX PR 05-SEP-2000; 2000DE-01044543.

XX PA (EPIG-) EPIGENOMICS AG.

XX PI Olek A, Piepenbrock C, Berlin K, Guetig D;

XX WPI; 2002-371829/40.

XX Determining the degree of cytosine methylation in genomic DNA, useful for  
PT diagnosis and prognosis, comprises selective hybridization of amplicons  
PT from chemically treated DNA.

XX PS Claim 12; 56pp + Sequence Listing; 56pp; German.

XX This invention describes a novel method for determining the degree of  
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
CC genomic sample of DNA. The sample is treated chemically to convert  
CC cytosine (C) but not methylated C, to uracil, then part of the genomic  
CC DNA that contains the target C is amplified to form a labeled amplicon.  
CC The amplicon is hybridised to two classes, each with at least one member,  
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the  
CC degree of hybridisation to both classes is determined from the label on  
CC the amplicon. From the ratio of labels hybridised to the two classes of  
CC oligomers, the degree of methylation is calculated. The method is used:  
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs  
CC and of a wide range of diseases, e.g. cancer, disorders of the central  
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,  
CC particularly by detecting mutations or single nucleotide polymorphisms  
CC (SNP's); and (ii) for differentiation of cell or tissue types and for  
CC investigating cell differentiation. The method allows the methylation  
CC status of many C residues to be determined simultaneously. ABQ13410-  
CC ABQ54121 represent genomic DNA sequences used to illustrate the method  
CC for determining the degree of cytosine methylation described in the  
CC disclosure of the invention

SQ Sequence 1234 BP; 380 A; 498 C; 151 G; 205 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 540 Length: 1234  
Score: 41.00 Matches: 6  
Percent Similarity: 100.00% Conservatives: 3  
Best Local Similarity: 66.67% Mismatches: 0  
Query Match: 82.00% Indels: 0  
DB: 6 Gaps: 0

US-10-014-101B-40 (1-9) x ABQ34201 (1-1234)

Oy 1 ValLeuGlyLeuGlyGlnPheCys 9

Db 703 GTGTCGGCGCGTGGTGGATTTCG 677

RESULT 39

AAS51508

ID AAS51508 standard; DNA; 1296 BP.

XX AC AAS51508;

XX AC

DT 13-FEB-2002 (first entry)

XX DE Pseudomonas aeruginosa DNA for cellular proliferation protein #93.

XX KW Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;

XX KW antibacterial; drug design.

XX OS Pseudomonas aeruginosa.

XX PN WO200170955-A2.

XX PD 27-SEP-2001.

XX PF 21-MAR-2001; 2001WO-US009180.

XX PR 21-MAR-2000; 2000US-0191078P.

XX PR 23-MAY-2000; 2000US-0206848P.

XX PR 26-MAY-2000; 2000US-0207727P.

XX PR 23-OCT-2000; 2000US-0242578P.

XX PR 27-NOV-2000; 2000US-0253625P.

XX PR 22-DEC-2000; 2000US-0257931P.

XX PR 18-FEB-2001; 2001US-0269308P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Haselbeck R, Ohlsen KL, Zyeckind JW, Wall D, Trawick JD, Carr GJ;

XX PI Yamamoto RT, Xu HH;

XX DR WPI; 2001-611495/70.

XX DR P-PSDB; AAU33649.

XX XX

XX New polynucleotides for the identification and development of

XX antibiotics, comprise sequences of antisense nucleic acids.

XX PS Claim 27; SEQ ID NO 4090; 51lpp; English.

XX CC The invention relates to antisense inhibitors of genes essential to

XX CC prokaryotic cellular proliferation, their use in identifying the genes,

XX CC their use in the discovery of novel antibiotics, the essential genes

XX CC themselves and the encoded proteins. The prokaryotes used are Escherichia

XX CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,

XX CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also

XX CC useful for the identification of potential new targets for antibiotic

XX CC development. The antisense nucleic acids can also be used to identify

XX CC proteins used in proliferation, to express these proteins, and to obtain

XX CC antibodies capable of binding to the expressed proteins. The proteins can

XX CC be used to screen compounds in rational drug discovery programmes. The

XX CC antisense nucleic acid sequence is also useful to screen for homologous

XX CC nucleic acids which are required for cell proliferation in a wide variety

XX CC of organisms. The present sequence encodes an essential prokaryotic

XX CC cellular proliferation protein. Note: The sequence data for this patent

CC did not form part of the printed specification, but was obtained in

CC electronic format directly from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 1296 BP; 172 A; 466 C; 384 G; 274 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 570 Length: 1296  
Score: 41.00 Matches: 7  
Percent Similarity: 88.89% Conservatives: 1  
Best Local Similarity: 77.78% Mismatches: 1  
Query Match: 82.00% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-40 (1-9) x AAS51508 (1-1296)

Oy 1 ValLeuGlyLeuGlyGlnPheCys 9

Db 1081 GTGATCGCGCGCTCGCGGCTTCTGC 1107

RESULT 40

ACA19484

ID ACA19484 standard; DNA; 1296 BP.

XX AC ACA19484;

XX AC

DT 19-JUN-2003 (first entry)

XX DE Prokaryotic essential gene #1141.

XX KW Antisense; ds; prokaryotic essential gene; cell proliferation;

XX KW drug design; gene.

XX OS Pseudomonas aeruginosa.

XX PN WO200277183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyeckind JW;

XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX DR WPI; 2003-029926/02.

XX DR P-PSDB; ABU15614.

XX XX

XX New antisense nucleic acids, useful for identifying proteins or screening

XX PT for homologous nucleic acids required for cellular proliferation to

XX PT isolate candidate molecules for rational drug discovery programs.

XX PS Claim 14; SEQ ID NO 7354; 1766pp; English.

XX CC The invention relates to an isolated nucleic acid comprising any one of

XX CC the 6213 antisense sequences given in the specification where expression

XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:

XX CC (1) a vector comprising a promoter operably linked to the nucleic acid

XX CC encoding a polypeptide whose expression is inhibited by the antisense

XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

XX CC polypeptide or its fragment whose expression is inhibited by the

XX CC antisense nucleic acid; (4) an antibody capable of specifically binding

XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

XX CC proliferation or the activity of a gene in an operon required for

XX CC proliferation; (7) identifying a compound that influences the activity of

XX CC the gene product or that has an activity against a biological pathway

CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
CC prokaryotic essential genes. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX  
SQ Sequence 1296 BP; 172 A; 466 C; 384 G; 274 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 570 Length: 1296  
Score: 41.00 Matches: 7  
Percent Similarity: 88.89% Conservative: 1  
Best Local Similarity: 77.78% Mismatches: 1  
Query Match: 82.00% Indels: 0  
DB: 8 Gaps: 0

US-10-014-101B-40 (1-9) x ACA19484 (1-1296)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9

Db 1081 GTGATCGGGCGCTCGGGCGCTTCTGC 1107

Search completed: February 18, 2005, 03:27:08  
Job time : 190.74 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 18, 2005, 02:33:38 ; Search time 53.1 Seconds  
(without alignments)  
277.335 Million cell updates/sec

Title: US-10-014-101B-40  
Perfect score: 50  
Sequence: 1 VLGLGQFC 9

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US10014101/runat.16022005.075846.7970/app.query.fasta.1.796  
-DB=Issued Patents NA -QFMT=faetap -SUFFIX=rni -MINMATCH=0.1 -LOOCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10014101 @CGN 1 1 124 @runat.16022005.075846.7970 -NCPUS=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	43	86.0	5183	1	US-08-459-568-3
C 2	43	86.0	5183	1	US-08-399-411-3
C 3	43	86.0	5868	3	US-08-516-859A-3
C 4	43	86.0	5868	3	US-09-586-472-3
C 5	43	86.0	5868	3	US-09-528-706-3
C 6	43	86.0	6173	4	US-09-949-016-5511
C 7	43	86.0	6173	4	US-09-949-016-5512
C 8	43	86.0	42672	4	US-09-949-016-17253
C 9	43	86.0	42672	4	US-09-949-016-17254
C 10	42	84.0	706	3	US-08-998-416-873
C 11	41	82.0	1296	4	US-09-252-991A-3696
C 12	41	82.0	1368	4	US-09-252-991A-3813

13	41	82.0	1605	3	US-09-124-541-3	Sequence 3, Appli
14	41	82.0	1605	4	US-09-663-326-3	Sequence 3, Appli
15	41	82.0	2784	4	US-09-252-991A-3769	Sequence 3769, Ap
16	41	82.0	6733	3	US-09-124-541-2	Sequence 2, Appli
17	41	82.0	6733	4	US-09-663-326-2	Sequence 2, Appli
18	41	82.0	8473	3	US-09-172-422-2	Sequence 2, Appli
19	41	82.0	8473	4	US-09-916-204-3	Sequence 3, Appli
C 19	41	82.0	48763	4	US-10-282-048-3	Sequence 3, Appli
C 20	41	82.0	48763	4	US-10-282-048-3	Sequence 3, Appli
C 21	41	82.0	174493	4	US-09-804-471A-3	Sequence 3, Appli
C 22	41	82.0	174493	4	US-10-238-709-3	Sequence 3, Appli
C 23	40	80.0	601	4	US-09-949-016-73087	Sequence 73087, A
C 24	40	80.0	2611	4	US-09-620-312D-925	Sequence 925, App
C 25	40	80.0	168174	4	US-10-071-411A-63	Sequence 63, Appli
C 26	40	80.0	168273	4	US-10-071-411A-2	Sequence 2, Appli
C 27	40	80.0	194937	4	US-09-949-016-17032	Sequence 17032, A
C 28	40	80.0	194937	4	US-09-949-016-17033	Sequence 17033, A
C 29	40	80.0	221545	4	US-09-949-016-13875	Sequence 13875, A
C 30	40	80.0	784019	4	US-09-949-016-14033	Sequence 14033, A
C 31	40	80.0	828152	4	US-09-949-016-12777	Sequence 12777, A
C 32	39	78.0	283	2	US-08-691-814B-88	Sequence 88, Appli
C 33	39	78.0	284	2	US-08-691-814B-91	Sequence 91, Appli
C 34	39	78.0	303	2	US-08-691-814B-90	Sequence 90, Appli
C 35	39	78.0	325	2	US-08-691-814B-89	Sequence 89, Appli
C 36	39	78.0	2549	3	US-09-467-082-3	Sequence 3, Appli
C 37	39	78.0	2549	4	US-09-394-455-5	Sequence 5, Appli
C 38	39	78.0	2549	4	US-09-949-016-3584	Sequence 3584, Ap
C 39	39	78.0	2608	4	US-09-394-455-35	Sequence 35, Appli
C 40	39	78.0	3846	2	US-08-691-814B-3	Sequence 3, Appli
C 41	39	78.0	3846	4	US-09-949-016-821	Sequence 821, App
C 42	39	78.0	3860	4	US-09-949-016-3964	Sequence 3964, Ap
C 43	39	78.0	9053	4	US-09-976-594-306	Sequence 306, App
C 44	39	78.0	12425	4	US-09-616-289-50	Sequence 50, Appli
C 45	39	78.0	15353	4	US-09-949-016-13454	Sequence 13454, A

ALIGNMENTS

RESULT 1  
US-08-459-568-3/c  
; Sequence 3, Application US/08459568  
; Patent No. 5811304  
; GENERAL INFORMATION:  
; APPLICANT: Huang, Shi  
; TITLE OF INVENTION: Retinoblastoma Protein - Interacting  
; TITLE OF INVENTION: Zinc Finger Proteins  
; NUMBER OF SEQUENCES: 93  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/459,568  
; FILING DATE: 02-JUN-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/399,411  
; FILING DATE: 06-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LJ 1264  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 3:

1

; Sequence 3, Application US/09586472  
; Patent No. 6323335  
; GENERAL INFORMATION:  
; APPLICANT: Huang, Shi  
; TITLE OF INVENTION: Retinoblastoma Protein - Interacting  
; Zinc Finger Proteins  
; NUMBER OF SEQUENCES: 106  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell & Flores LLP  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/586,472  
; FILING DATE: 01-Jun-2000  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/528,706  
; FILING DATE: 17-MAR-2000  
; APPLICATION NUMBER: US 08/516,859  
; FILING DATE: 18-AUG-1995  
; APPLICATION NUMBER: US 08/399,411  
; FILING DATE: 06-MAR-1995  
; APPLICATION NUMBER: US 08/292,683  
; FILING DATE: 18-AUG-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LJ 4130  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5868 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 121..5278  
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-09-586-472-3

Alignment Scores:  
Pred. No.: 178 Length: 5868  
Score: 43.00 Matches: 8  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 1  
Query Match: 86.00% Indels: 0  
DB: 3 Gaps: 0

US-10-014-101B-40 (1-9) x US-09-586-472-3 (1-5868)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
Db 1728 GTTCTGGTGGCGCTGGCGCTGTTCTGC 1702  
RESULT 5  
US-09-528-706-3/c  
; Sequence 3, Application US/09528706  
; Patent No. 6468985  
; GENERAL INFORMATION:  
; APPLICANT: Huang, Shi  
; TITLE OF INVENTION: Retinoblastoma Protein - Interacting  
; Zinc Finger Proteins

; NUMBER OF SEQUENCES: 106  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell & Flores LLP  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/528,706  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/516,859  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/292,683  
; FILING DATE: 18-AUG-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LJ 1776  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5868 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 121..5278  
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-09-528-706-3

Alignment Scores:  
Pred. No.: 178 Length: 5868  
Score: 43.00 Matches: 8  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 1  
Query Match: 86.00% Indels: 0  
DB: 3 Gaps: 0

US-10-014-101B-40 (1-9) x US-09-528-706-3 (1-5868)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
Db 1728 GTTCTGGTGGCGCTGGCGCTGTTCTGC 1702

RESULT 6

US-09-949-016-5511/c  
; Sequence 5511, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5511

; LENGTH: 6173

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-5511

Alignment Scores:

Pred. No.: 190 Length: 6173  
Score: 43.00 Matches: 8  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 1  
Query Match: 86.00% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-40 (1-9) x US-09-949-016-5511 (1-6173)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9

Db 1182 GTTCTGGGTGGCCTGGCCTGTTCTGC 1156

RESULT 7

US-09-949-016-5512/c

; Sequence 5512, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949, 016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5512

; LENGTH: 6173

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-5512

Alignment Scores:

Pred. No.: 190 Length: 6173  
Score: 43.00 Matches: 8  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 1  
Query Match: 86.00% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-40 (1-9) x US-09-949-016-5512 (1-6173)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9

Db 1182 GTTCTGGGTGGCCTGGCCTGTTCTGC 1156

RESULT 8

US-09-949-016-17253/c

; Sequence 17253, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949, 016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 17253

; LENGTH: 42672

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-17253

Alignment Scores:

Pred. No.: 1.89e+03 Length: 42672  
Score: 43.00 Matches: 8  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 1  
Query Match: 86.00% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-40 (1-9) x US-09-949-016-17253 (1-42672)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9

Db 31998 GTTCTGGGTGGCCTGGCCTGTTCTGC 31972

RESULT 9

US-09-949-016-17254/c

; Sequence 17254, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949, 016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 17254

; LENGTH: 42672

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-17254

Alignment Scores:

Pred. No.: 1.89e+03 Length: 42672  
Score: 43.00 Matches: 8  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 1  
Query Match: 86.00% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-40 (1-9) x US-09-949-016-17254 (1-42672)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9

Db 31998 GTTCTGGGTGGCCTGGCCTGTTCTGC 31972

RESULT 10

US-08-998-416-873

; Sequence 873, Application US/08998416

; Patent No. 6239264

; GENERAL INFORMATION:

; APPLICANT: Philippaen, Peter

; APPLICANT: Pohlmann, Rainer

; APPLICANT: Steiner, Sabine

APPLICANT: Mohr, Christine  
APPLICANT: Wendland, Jorgen  
APPLICANT: Knechtie, Philipp  
APPLICANT: Rebeschung, Corinne  
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPHII  
TITLE OF INVENTION: AND USES THEREOF  
NUMBER OF SEQUENCES: 1152  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 6239264artis Corporation  
STREET: 3054 Cornwallis Road  
CITY: Research Triangle Park  
STATE: No. 6239264th Carolina  
COUNTRY: USA  
ZIP: 27709

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/998,416  
FILING DATE: 24-DEC-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: CH 0016/97  
FILING DATE: 31-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Meigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8587  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 873:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 706 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: PAG1548RP  
US-08-998-416-873

Alignment Scores:  
Pred. No.: 22.8 Length: 706  
Score: 42.00 Matches: 7  
Percent Similarity: 87.50% Conservative: 0  
Best Local Similarity: 87.50% Mismatches: 1  
Query Match: 84.00% Indels: 0  
DB: 3 Gaps: 0

US-10-014-101B-40 (1-9) x US-08-998-416-873 (1-706)

QY 2 LeuGlyGlyLeuGlyGlnPheCys 9  
DB 344 CTGGCGGGTTGGCAGATTCTGC 367

## RESULT 11

US-09-252-991A-3696/c  
Sequence 3696, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR FILING DATE: 1998-02-18  
PRIOR FILING DATE: 1998-02-18  
PRIOR FILING DATE: 1998-02-18  
PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 3696  
LENGTH: 1296  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-3696

Alignment Scores:  
Pred. No.: 74.3 Length: 1296  
Score: 41.00 Matches: 7  
Percent Similarity: 88.89% Conservative: 1  
Best Local Similarity: 77.78% Mismatches: 1  
Query Match: 82.00% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-40 (1-9) x US-09-252-991A-3696 (1-1296)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
DB 225 GTGATCGCGGCTCGCGGCTTCTGC 199

## RESULT 12

US-09-252-991A-3813  
Sequence 3813, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 3813  
LENGTH: 1368  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-3813

Alignment Scores:  
Pred. No.: 79.3 Length: 1368  
Score: 41.00 Matches: 7  
Percent Similarity: 88.89% Conservative: 1  
Best Local Similarity: 77.78% Mismatches: 1  
Query Match: 82.00% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-40 (1-9) x US-09-252-991A-3813 (1-1368)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
DB 1153 GTGATCGCGGCTCGCGGCTTCTGC 1179

## RESULT 13

US-09-124-541-3  
Sequence 3, Application US/09124541A  
Patent No. 6229066  
GENERAL INFORMATION:  
APPLICANT: Morris Ph.D., Roy O.  
TITLE OF INVENTION: A CYTOKININ OXIDASE  
FILE REFERENCE: UM01490  
CURRENT APPLICATION NUMBER: US/09/124,541A  
CURRENT FILING DATE: 1998-07-29  
EARLIER FILING DATE: 1997-07-30  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3  
LENGTH: 1605

; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1605)  
US-09-124-541-3

Alignment Scores:  
Pred. No.: 95.9 Length: 1605  
Score: 41.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 82.00% Indels: 0  
DB: 3 Gaps: 0

US-10-014-101B-40 (1-9) x US-09-124-541-3 (1-1605)

QY 1 ValLeuGlyLeuGlyGlnPhe 8  
Db 673 GTCTGGGCGGCTGGGGCAGTTC 696

## RESULT 14

US-09-663-326-3  
; Sequence 3, Application US/09663326  
; Patent No. 6617497  
; GENERAL INFORMATION:  
; APPLICANT: Morris Ph.D., Roy O.  
; TITLE OF INVENTION: A CYTOKININ OXIDASE  
; FILE REFERENCE: UM01490  
; CURRENT APPLICATION NUMBER: US/09/663,326  
; CURRENT FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: 60/054,268  
; PRIOR FILING DATE: 1997-07-30  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1605  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1605)  
US-09-663-326-3

Alignment Scores:  
Pred. No.: 95.9 Length: 1605  
Score: 41.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 82.00% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-40 (1-9) x US-09-663-326-3 (1-1605)

QY 1 ValLeuGlyLeuGlyGlnPhe 8  
Db 673 GTCTGGGCGGCTGGGGCAGTTC 696

## RESULT 15

US-09-252-991A-3769  
; Sequence 3769, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 3769  
; LENGTH: 2784  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-3769

Alignment Scores:  
Pred. No.: 184 Length: 2784  
Score: 41.00 Matches: 7  
Percent Similarity: 88.89% Conservative: 1  
Best Local Similarity: 77.78% Mismatches: 1  
Query Match: 82.00% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-40 (1-9) x US-09-252-991A-3769 (1-2784)

QY 1 ValLeuGlyLeuGlyGlnPheCys 9  
Db 1118 GTGATCGGGCGCTCGGGCTTCTGC 1144

## RESULT 16

US-09-124-541-2  
; Sequence 2, Application US/09124541A  
; Patent No. 6229066  
; GENERAL INFORMATION:  
; APPLICANT: Morris Ph.D., Roy O.  
; TITLE OF INVENTION: A CYTOKININ OXIDASE  
; FILE REFERENCE: UM01490  
; CURRENT APPLICATION NUMBER: US/09/124,541A  
; CURRENT FILING DATE: 1998-07-29  
; EARLIER APPLICATION NUMBER: 60/054,268  
; EARLIER FILING DATE: 1997-07-30  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 6733  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: gene  
; LOCATION: (1)..(6733)  
; OTHER INFORMATION: genomic sequence for a cytokinin oxidase from Zea  
; OTHER INFORMATION: mays  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1497)..(2111)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (2524)..(3216)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (3311)..(3607)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (5697)  
US-09-124-541-2

Alignment Scores:  
Pred. No.: 527 Length: 6733  
Score: 41.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 82.00% Indels: 0  
DB: 3 Gaps: 0

US-10-014-101B-40 (1-9) x US-09-124-541-2 (1-6733)

QY 1 ValLeuGlyLeuGlyGlnPhe 8  
Db 2581 GTCTGGGCGGCTGGGGCAGTTC 2604

## RESULT 17

US-09-663-326-2  
; Sequence 2, Application US/09663326  
; Patent No. 6617497  
; GENERAL INFORMATION:  
; APPLICANT: Morris Ph.D., Roy O.  
; TITLE OF INVENTION: A CYTOKININ OXIDASE  
; FILE REFERENCE: UM01490  
; CURRENT APPLICATION NUMBER: US/09/663,326  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: 60/054,268  
; PRIOR FILING DATE: 1997-07-30  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 6733  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: gene  
; LOCATION: (1)..(6733)  
; OTHER INFORMATION: genomic sequence for a cytokinin oxidase from Zea  
; OTHER INFORMATION: mays  
; NAME/KEY: CDS  
; LOCATION: (1497)..(2111)  
; NAME/KEY: CDS  
; LOCATION: (2524)..(3216)  
; NAME/KEY: CDS  
; LOCATION: (3311)..(3607)  
; NAME/KEY: unsure  
; LOCATION: (5697)  
US-09-663-326-2  
Alignment Scores:  
Pred. No.: 527 Length: 6733  
Score: 41.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 82.00% Indels: 0  
DB: 4 Gaps: 0  
US-10-014-101B-40 (1-9) x US-09-663-326-2 (1-6733)  
QY 1 ValLeuGlyLeuGlyGlnPhe 8  
DB 2581 GTCTGGCGGCTGGGGCAGTTC 2604  
RESULT 18  
US-09-172-422-2  
; Sequence 2, Application US/09172422A  
; Patent No. 6300485  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Arwen E.  
; APPLICANT: Chiu, Choi Ying  
; APPLICANT: Duhl, David  
; APPLICANT: Gorman, Susan W.  
; APPLICANT: Leng, Song  
; APPLICANT: Sheffield, Val  
; APPLICANT: Welch, Juliet  
; TITLE OF INVENTION: MYOSIN IXA AND CYCLIC NUCLEOTIDE GATED  
; TITLE OF INVENTION: CHANNEL-15 (CMGC-15) POLYNUCLEOTIDES, POLYPEPTIDES,  
; TITLE OF INVENTION: COMPOSITIONS, METHODS, AND USES THEREOF  
; FILE REFERENCE: 200130.442  
; CURRENT APPLICATION NUMBER: US/09/172,422A  
; CURRENT FILING DATE: 1998-10-14  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 8473  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-172-422-2  
Alignment Scores:  
Pred. No.: 527 Length: 8473  
Score: 41.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 3  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 82.00% Indels: 0  
DB: 4 Gaps: 0

Pred. No.: 692 Length: 8473  
Score: 41.00 Matches: 7  
Percent Similarity: 87.50% Conservative: 0  
Best Local Similarity: 87.50% Mismatches: 1  
Query Match: 82.00% Indels: 0  
DB: 3 Gaps: 0  
US-10-014-101B-40 (1-9) x US-09-172-422-2 (1-8473)  
QY 2 LeuGlyGlyLeuGlyGlnPheCys 9  
DB 59 CTCGGTGGCTCGGCTCGTTCTGC 82  
RESULT 19  
US-09-916-204-3/c  
; Sequence 3, Application US/09916204  
; Patent No. 6638745  
; GENERAL INFORMATION:  
; APPLICANT: Wei, Ming-Hui et al.  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001164CIP  
; CURRENT APPLICATION NUMBER: US/09/916,204  
; CURRENT FILING DATE: 2001-07-24  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 48763  
; TYPE: DNA  
; ORGANISM: Human  
US-09-916-204-3  
Alignment Scores:  
Pred. No.: 5.54e+03 Length: 48763  
Score: 41.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 3  
Best Local Similarity: 66.67% Mismatches: 0  
Query Match: 82.00% Indels: 0  
DB: 4 Gaps: 0  
US-10-014-101B-40 (1-9) x US-09-916-204-3 (1-48763)  
QY 1 ValLeuGlyLeuGlyGlnPheCys 9  
DB 23030 GTTCTTGGGGAATGGGAGGTATTGT 23004  
RESULT 20  
US-10-282-048-3/c  
; Sequence 3, Application US/10282048  
; Patent No. 6692948  
; GENERAL INFORMATION:  
; APPLICANT: Wei, Ming-Hui et al.  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001164CIP-DIV  
; CURRENT APPLICATION NUMBER: US/10/282,048  
; CURRENT FILING DATE: 2002-10-29  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 48763  
; TYPE: DNA  
; ORGANISM: Human  
US-10-282-048-3  
Alignment Scores:  
Pred. No.: 5.54e+03 Length: 48763  
Score: 41.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 3  
Best Local Similarity: 66.67% Mismatches: 0  
Query Match: 82.00% Indels: 0

```
DB: 4 Gaps: 0
US-10-014-101B-40 (1-9) x US-10-282-048-3 (1-48763)
Qy 1 ValLeuGlyGlyLeuGlyGlnPheCys 9
Db 23030 GTTCTTGGGGAATGGGAGGTATTGT 23004

RESULT 21
US-09-804-471A-3/c
; Sequence 3, Application US/09804471A
; Patent No. 6479269
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001164
; CURRENT APPLICATION NUMBER: US/09/804,471A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 174493
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(174493)
; OTHER INFORMATION: n = A,T,C or G
US-09-804-471A-3

Alignment Scores:
Pred. No.: 2,51e+04 Length: 174493
Score: 41.00 Matches: 6
Percent Similarity: 100.00% Conservative: 3
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 82.00% Indels: 0
DB: 4 Gaps: 0

US-10-014-101B-40 (1-9) x US-09-804-471A-3 (1-174493)
Qy 1 ValLeuGlyGlyLeuGlyGlnPheCys 9
Db 99462 GTTCTTGGGGAATGGGAGGTATTGT 99436

RESULT 22
US-10-238-709-3/c
; Sequence 3, Application US/10238709
; Patent No. 6680188
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001164DIV
; CURRENT APPLICATION NUMBER: US/10/238,709
; CURRENT FILING DATE: 2002-09-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 174493
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(174493)
; OTHER INFORMATION: n = A,T,C or G
US-10-238-709-3

Alignment Scores:
Pred. No.: 2,51e+04 Length: 174493
Score: 41.00 Matches: 6
Percent Similarity: 100.00% Conservative: 3
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 82.00% Indels: 0
DB: 4 Gaps: 0

US-10-014-101B-40 (1-9) x US-09-804-471A-3 (1-174493)
Qy 1 ValLeuGlyGlyLeuGlyGlnPheCys 9
Db 99462 GTTCTTGGGGAATGGGAGGTATTGT 99436

RESULT 23
US-09-949-016-73087/c
; Sequence 73087, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73087
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-73087

Alignment Scores:
Pred. No.: 47.2 Length: 601
Score: 40.00 Matches: 7
Percent Similarity: 77.78% Conservative: 0
Best Local Similarity: 77.78% Mismatches: 2
Query Match: 80.00% Indels: 0
DB: 4 Gaps: 0

US-10-014-101B-40 (1-9) x US-09-949-016-73087 (1-601)
Qy 1 ValLeuGlyGlyLeuGlyGlnPheCys 9
Db 64 GTGCAGGAGGTTTGGGTTCCTCTGTC 38

RESULT 24
US-09-620-312D-925
; Sequence 925, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
```

; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 784CIP2B  
; CURRENT APPLICATION NUMBER: US/09/620,312D  
; CURRENT FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1105  
; SOFTWARE: pt\_FL\_genes Version 1.0  
; SEQ ID NO 925  
; LENGTH: 2611  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (290)..(1885)  
US-09-620-312D-925

Alignment Scores:  
Pred. No.: 271 Length: 2611  
Score: 40.00 Matches: 7  
Percent Similarity: 88.89% Conservative: 1  
Best Local Similarity: 77.78% Mismatches: 1  
Query Match: 80.00% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-40 (1-9) x US-09-620-312D-925 (1-2611)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
|||||  
DB 101 GTCTCGCGCGCTGGTGGCTACTGTC 127

## RESULT 25

US-10-071-411A-63/c  
; Sequence 63, Application US/10071411A  
; Patent No. 6797475  
; APPLICANT: Glenn Barnes  
; APPLICANT: Joanne Meyer  
; TITLE OF INVENTION: Detection of Polymorphisms in the Human  
; TITLE OF INVENTION: 5-Lipoxygenase Gene  
; FILE REFERENCE: MRI-021  
; CURRENT APPLICATION NUMBER: US/10/071,411A  
; CURRENT FILING DATE: 2002-02-07  
; PRIOR APPLICATION NUMBER: 60/267,515  
; PRIOR FILING DATE: 2001-02-08  
; PRIOR APPLICATION NUMBER: 60/314,248  
; PRIOR FILING DATE: 2001-08-21  
; NUMBER OF SEQ ID NOS: 66  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 63  
; LENGTH: 168174  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(168174)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-071-411A-63

Alignment Scores:  
Pred. No.: 3,79e+04 Length: 168174  
Score: 40.00 Matches: 7  
Percent Similarity: 88.89% Conservative: 1  
Best Local Similarity: 77.78% Mismatches: 1  
Query Match: 80.00% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-40 (1-9) x US-10-071-411A-63 (1-168174)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
|||||

DB 22730 GTTCTGGCAGGCTGGGGAGATTCTGC 22704

## RESULT 26

US-10-071-411A-2/c  
; Sequence 2, Application US/10071411A  
; Patent No. 6797475  
; GENERAL INFORMATION:  
; APPLICANT: Glenn Barnes  
; APPLICANT: Joanne Meyer  
; TITLE OF INVENTION: Detection of Polymorphisms in the Human  
; TITLE OF INVENTION: 5-Lipoxygenase Gene  
; FILE REFERENCE: MRI-021  
; CURRENT APPLICATION NUMBER: US/10/071,411A  
; CURRENT FILING DATE: 2002-02-07  
; PRIOR APPLICATION NUMBER: 60/267,515  
; PRIOR FILING DATE: 2001-02-08  
; PRIOR APPLICATION NUMBER: 60/314,248  
; PRIOR FILING DATE: 2001-08-21  
; NUMBER OF SEQ ID NOS: 66  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 168273  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(168273)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-071-411A-2

Alignment Scores:  
Pred. No.: 3,79e+04 Length: 168273  
Score: 40.00 Matches: 7  
Percent Similarity: 88.89% Conservative: 1  
Best Local Similarity: 77.78% Mismatches: 1  
Query Match: 80.00% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-40 (1-9) x US-10-071-411A-2 (1-168273)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
|||||

DB 22780 GTTCTGGCAGGCTGGGGAGATTCTGC 22754

## RESULT 27

US-09-949-016-17032  
; Sequence 17032, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17032  
; LENGTH: 194937  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-17032

Alignment Scores:  
Pred. No.: 4,51e+04 Length: 194937  
Score: 40.00 Matches: 7  
Percent Similarity: 88.89% Conservative: 1

Best Local Similarity: 77.78% Mismatches: 1  
Query Match: 80.00% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-40 (1-9) x US-09-949-016-17032 (1-194937)

Qy 1 ValLeuGlyLeuGlyGlnPheCys 9

Db 28384 GTCTAGGAGGAGTGGGAGGTTTGT 28410

## RESULT 28

US-09-949-016-17033  
; Sequence 17033, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17033  
; LENGTH: 194937  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-17033

Alignment Scores:  
Pred. No.: 4.51e+04 Length: 194937  
Score: 40.00 Matches: 7  
Percent Similarity: 88.89% Conservative: 1  
Best Local Similarity: 77.78% Mismatches: 1  
Query Match: 80.00% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-40 (1-9) x US-09-949-016-17033 (1-194937)

Qy 1 ValLeuGlyLeuGlyGlnPheCys 9

Db 28384 GTCTAGGAGGAGTGGGAGGTTTGT 28410

## RESULT 29

US-09-949-016-13875  
; Sequence 13875, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13875  
; LENGTH: 221545  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-13875

Alignment Scores:  
Pred. No.: 5.24e+04 Length: 221545  
Score: 40.00 Matches: 7  
Percent Similarity: 77.78% Conservative: 0  
Best Local Similarity: 77.78% Mismatches: 2  
Query Match: 80.00% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-40 (1-9) x US-09-949-016-13875 (1-221545)

Qy 1 ValLeuGlyLeuGlyGlnPheCys 9

Db 198801 GTGGCAGGAGTTGGGTTCTTCTGC 198827

## RESULT 30

US-09-949-016-14033  
; Sequence 14033, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14033  
; LENGTH: 784019  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(784019)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-14033

Alignment Scores:  
Pred. No.: 2.27e+05 Length: 784019  
Score: 40.00 Matches: 7  
Percent Similarity: 87.50% Conservative: 0  
Best Local Similarity: 87.50% Mismatches: 1  
Query Match: 80.00% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-40 (1-9) x US-09-949-016-14033 (1-784019)

Qy 2 LeuGlyLeuGlyGlnPheCys 9

Db 254504 CTGGGAGGAGTGGGACAGATCTGT 254527

## RESULT 31

US-09-949-016-12777  
; Sequence 12777, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03

;; PRIOR APPLICATION NUMBER: 60/231,498  
;; PRIOR FILING DATE: 2000-09-08  
;; NUMBER OF SEQ ID NOS: 207012  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 12777  
;; LENGTH: 828152  
;; TYPE: DNA  
;; ORGANISM: Human  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; LOCATION: (1)...(828152)  
;; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-12777

Alignment Scores: 2.41e+05 Length: 828152  
Pred. No.: 40.00 Matches: 7  
Score: 40.00  
Percent Similarity: 87.50% Conservative: 0  
Best Local Similarity: 87.50% Mismatches: 1  
Query Match: 80.00% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-40 (1-9) x US-09-949-016-12777 (1-828152)

QY 2 LeuGlyGlyLeuGlyGlnPheCys 9

Db 250637 CTGGGAGGACTGGGACAGATCTGT 250660

RESULT 32

US-08-691-814B-88  
; Sequence 88, Application US/08691814B  
; Patent No. 5981218  
; GENERAL INFORMATION:  
; APPLICANT: Rio, Marie-Christine  
; APPLICANT: Tomasetto, Catherine  
; APPLICANT: Bassett, Paul  
; APPLICANT: Byrne, Jennifer  
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Useful  
; TITLE OF INVENTION: as Leukemia Markers and in Breast Cancer Prognosis  
; NUMBER OF SEQUENCES: 124  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
; STREET: 1100 New York Ave, NW, Suite 600  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/691,814B  
; FILING DATE: 31-JUL-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/002,183  
; FILING DATE: 09-AUG-1995  
; NAME: Steffe, Eric K.  
; REGISTRATION NUMBER: 36,688  
; REFERENCE/DOCKET NUMBER: 1383.0090001  
; TELEPHONE: 202-371-2600  
; TELEFAX: 202-371-2543  
; INFORMATION FOR SEQ ID NO: 88:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 283 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: both  
; MOLECULE TYPE: cdna

US-08-691-814B-88

Alignment Scores: 30.5 Length: 283  
Pred. No.: 39.00 Matches: 7  
Score: 39.00  
Percent Similarity: 77.78% Conservative: 0  
Best Local Similarity: 77.78% Mismatches: 2  
Query Match: 78.00% Indels: 0  
DB: 2 Gaps: 0

US-10-014-101B-40 (1-9) x US-08-691-814B-88 (1-283)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9

Db 38 GTCTTGGTGGCTGGTGCATCTGT 64

RESULT 33

US-08-691-814B-91  
; Sequence 91, Application US/08691814B  
; Patent No. 5981218  
; GENERAL INFORMATION:  
; APPLICANT: Rio, Marie-Christine  
; APPLICANT: Tomasetto, Catherine  
; APPLICANT: Bassett, Paul  
; APPLICANT: Byrne, Jennifer  
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Useful  
; TITLE OF INVENTION: as Leukemia Markers and in Breast Cancer Prognosis  
; NUMBER OF SEQUENCES: 124  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
; STREET: 1100 New York Ave, NW, Suite 600  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/691,814B  
; FILING DATE: 31-JUL-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/002,183  
; FILING DATE: 09-AUG-1995  
; NAME: Steffe, Eric K.  
; REGISTRATION NUMBER: 36,688  
; REFERENCE/DOCKET NUMBER: 1383.0090001  
; TELEPHONE: 202-371-2600  
; TELEFAX: 202-371-2543  
; INFORMATION FOR SEQ ID NO: 91:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 284 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: both  
; MOLECULE TYPE: cdna

Alignment Scores: 30.7 Length: 284  
Pred. No.: 39.00 Matches: 7  
Score: 39.00  
Percent Similarity: 77.78% Conservative: 0  
Best Local Similarity: 77.78% Mismatches: 2  
Query Match: 78.00% Indels: 0  
DB: 2 Gaps: 0

US-10-014-101B-40 (1-9) x US-08-691-814B-91 (1-284)

```

; APPLICANT: Byrne, Jennifer
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Useful
; TITLE OF INVENTION: as Leukemia Markers and in Breast Cancer Prognosis
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Ave, NW, Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/691,814B
; FILING DATE: 31-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/002,183
; FILING DATE: 09-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Steffe, Eric K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1383.0090001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2543
; INFORMATION FOR SEQ ID NO: 89:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 325 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; US-08-691-814B-89

Alignment Scores:
Pred. No.: 36 Length: 325
Score: 39.00 Matches: 7
Percent Similarity: 77.78% Conservative: 0
Best Local Similarity: 77.78% Mismatches: 2
Query Match: 78.00% Indels: 0
DB: 2 Gaps: 0

US-10-014-101B-40 (1-9) x US-08-691-814B-89 (1-325)

QY 1 ValLeuglyGlyLeuglyGlnPheCys 9
Db 1 GTCTTGGGTGGCGCTGTCGATTCTGT 27

RESULT 36
US-09-467-082-3/c
; Sequence 3, Application US/09467082
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Lex M. Cowest
; TITLE OF INVENTION: ANTISENSE MODULATION OF PKA CATALYTIC SUBUNIT C-ALPHA EXPRESSION
; FILE REFERENCE: RTS-0088
; CURRENT APPLICATION NUMBER: US/09/467,082
; CURRENT FILING DATE: 1999-12-17
; NUMBER OF SEQ ID NOS: 49
; SEQ ID NO 3
; LENGTH: 2549
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (81)..(1136)
; NAME/KEY: unsure
; LOCATION: 6

```

; OTHER INFORMATION: unknown  
US-09-467-082-3

Alignment Scores:                    Length:                    2549  
Pred. No.:                    416  
Score:                    39.00  
Percent Similarity:                    87.50%  
Best Local Similarity:                    87.50%  
Query Match:                    78.00%  
Indels:                    0  
DB:                    3

US-10-014-101B-40 (1-9) x US-09-467-082-3 (1-2549)

QY                    2 LeuGlyGlyLeuGlyGlnPheCys 9  
                  |||||  
Db                    2169 TTGGGGGACTTGGCCTCTTCTGT 2146

RESULT 37

US-09-394-455-5/c  
; Sequence 5, Application US/09394455  
; Patent No. 6531305  
; GENERAL INFORMATION:  
; APPLICANT: Witman, George F.  
; APPLICANT: San Agustin, Jovenal  
; APPLICANT: Leszyk, John D.  
; TITLE OF INVENTION: SPERM ASSOCIATED PROTEIN KINASE POLYPEPTIDES, CORRESPONDING  
; FILE REFERENCE: 07917/078001  
; CURRENT APPLICATION NUMBER: US/09/394,455  
; CURRENT FILING DATE: 1999-09-10  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: US 60/099,771  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 2549  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (81)...(1133)  
; NAME/KEY: misc feature  
; LOCATION: (1)...(2549)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-394-455-5

Alignment Scores:                    Length:                    2549  
Pred. No.:                    416  
Score:                    39.00  
Percent Similarity:                    87.50%  
Best Local Similarity:                    87.50%  
Query Match:                    78.00%  
Indels:                    0  
DB:                    4

US-10-014-101B-40 (1-9) x US-09-394-455-5 (1-2549)

QY                    2 LeuGlyGlyLeuGlyGlnPheCys 9  
                  |||||  
Db                    2169 TTGGGGGACTTGGCCTCTTCTGT 2146

RESULT 38

US-09-949-016-3584/c  
; Sequence 3584, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3584  
; LENGTH: 2549  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-3584

Alignment Scores:                    Length:                    2549  
Pred. No.:                    416  
Score:                    39.00  
Percent Similarity:                    87.50%  
Best Local Similarity:                    87.50%  
Query Match:                    78.00%  
Indels:                    0  
DB:                    4

US-10-014-101B-40 (1-9) x US-09-949-016-3584 (1-2549)

QY                    2 LeuGlyGlyLeuGlyGlnPheCys 9  
                  |||||  
Db                    2168 TTGGGGGACTTGGCCTCTTCTGT 2145

RESULT 39

US-09-394-455-35/c  
; Sequence 35, Application US/09394455  
; Patent No. 6531305  
; GENERAL INFORMATION:  
; APPLICANT: Witman, George F.  
; APPLICANT: San Agustin, Jovenal  
; APPLICANT: Leszyk, John D.  
; TITLE OF INVENTION: SPERM ASSOCIATED PROTEIN KINASE POLYPEPTIDES, CORRESPONDING  
; FILE REFERENCE: 07917/078001  
; CURRENT APPLICATION NUMBER: US/09/394,455  
; CURRENT FILING DATE: 1999-09-10  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: US 60/099,771  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 35  
; LENGTH: 2608  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (164)...(1192)  
US-09-394-455-35

Alignment Scores:                    Length:                    2608  
Pred. No.:                    428  
Score:                    39.00  
Percent Similarity:                    87.50%  
Best Local Similarity:                    87.50%  
Query Match:                    78.00%  
Indels:                    0  
DB:                    4

US-10-014-101B-40 (1-9) x US-09-394-455-35 (1-2608)

QY                    2 LeuGlyGlyLeuGlyGlnPheCys 9  
                  |||||  
Db                    2228 TTGGGGGACTTGGCCTCTTCTGT 2205

RESULT 40

US-08-691-814B-3  
; Sequence 3, Application US/08691814B  
; Patent No. 5981218  
; GENERAL INFORMATION:  
; APPLICANT: Rio, Marie-Christine  
; APPLICANT: Tomasetto, Catherine

APPLICANT: Basset, Paul  
APPLICANT: Byrne, Jennifer  
TITLE OF INVENTION: Isolated Nucleic Acid Molecules Useful  
TITLE OF INVENTION: as Leukemia Markers and in Breast Cancer Prognosis  
NUMBER OF SEQUENCES: 124  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
STREET: 1100 New York Ave, NW, Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/691,814B  
FILING DATE: 31-JUL-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/002,183  
FILING DATE: 09-AUG-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Steffe, Eric K.  
REGISTRATION NUMBER: 36,688  
REFERENCE/DOCKET NUMBER: 1383.0090001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2543  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3846 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: both  
MOLECULE TYPE: cdna  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 76..858  
US-08-691-814B-3

Alignment Scores:  
Pred. No.: 679 Length: 3846  
Score: 39.00 Matches: 7  
Percent Similarity: 77.78% Conservative: 0  
Best Local Similarity: 77.78% Mismatches: 2  
Query Match: 78.00% Indels: 0  
DB: 2 Gaps: 0

US-10-014-101B-40 (1-9) x US-08-691-814B-3 (1-3846)

Qy 1 valLeuGlyGlyLeuGlnPheCys 9  
Db 2900 GTCTTGGTGGCCCTGGTGCATTCTGT 2926

Search completed: February 18, 2005, 07:19:02  
Job time : 224.35 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: February 18, 2005, 05:26:14 ; Search time 424.26 Seconds  
(without alignments)  
125.382 Million cell updates/sec

Title: US-10-014-101B-40  
Perfect score: 50  
Sequence: 1 VLGGIGQFC 9

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 5384158 seqs, 2955248155 residues

Total number of hits satisfying chosen parameters: 10768316

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10014101/runat\_16022005\_075848\_8061/app\_query.fasta\_1.796  
-DB=Published Applications NA\_QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62  
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=40 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10014101 @CGN 1 659 @runat\_16022005\_075848\_8061  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
20: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
21: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
22: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	45	90.0	65047	13	US-10-087-192-259	Sequence 259, App
2	43	86.0	503	16	US-10-029-386-9173	Sequence 9173, Ap
3	43	86.0	789	16	US-10-029-386-22854	Sequence 22854, A
c 4	43	86.0	5868	14	US-10-024-450-3	Sequence 3, Appli
5	43	86.0	6113	15	US-10-311-455-1795	Sequence 1795, Ap
6	43	86.0	6436	15	US-10-311-455-654	Sequence 654, App
7	43	86.0	7603	15	US-10-311-455-281	Sequence 281, App
8	43	86.0	7603	16	US-10-240-452-29	Sequence 29, Appl
9	43	86.0	9504	15	US-10-240-453-281	Sequence 281, App
10	43	86.0	254087	13	US-10-087-192-223	Sequence 223, App
11	42	84.0	556	18	US-10-425-115-102522	Sequence 102522,
c 12	42	84.0	564	13	US-10-027-632-280541	Sequence 280541,
c 13	42	84.0	564	17	US-10-027-632-280541	Sequence 280541,
c 14	42	84.0	1230	17	US-10-424-599-70249	Sequence 70249, A
c 15	42	84.0	2139	18	US-10-437-963-82245	Sequence 82245, A
16	42	84.0	7329	17	US-10-221-714A-396	Sequence 396, App
17	42	84.0	13321	17	US-10-221-714A-143	Sequence 143, App
18	42	84.0	3673778	16	US-10-312-841-1	Sequence 1, Appli
19	41	82.0	201	18	US-10-741-601-10567	Sequence 10567, A
20	41	82.0	201	19	US-10-741-600-26062	Sequence 26062, A
c 21	41	82.0	255	10	US-09-930-213-609	Sequence 609, App
22	41	82.0	369	18	US-10-425-115-74928	Sequence 74928, A
c 23	41	82.0	537	17	US-10-282-122A-27902	Sequence 27902, A
24	41	82.0	552	16	US-10-326-184-15	Sequence 15, Appl
25	41	82.0	566	16	US-10-326-184-14	Sequence 14, Appl
26	41	82.0	582	16	US-10-326-184-18	Sequence 18, Appl
27	41	82.0	588	18	US-10-425-115-94827	Sequence 94827, A
28	41	82.0	635	16	US-10-326-184-16	Sequence 16, Appl
29	41	82.0	1088	15	US-10-106-698-1224	Sequence 1224, Ap
30	41	82.0	1141	17	US-10-425-114-10886	Sequence 10886, A
31	41	82.0	1142	17	US-10-424-599-1725	Sequence 1725, Ap
32	41	82.0	1234	18	US-10-363-345A-20791	Sequence 20791, A
c 33	41	82.0	1234	18	US-10-363-345A-20792	Sequence 20792, A
34	41	82.0	1238	17	US-10-437-963-81681	Sequence 81681, A
35	41	82.0	1257	18	US-10-425-114-3248	Sequence 3248, Ap
36	41	82.0	1358	18	US-10-425-115-150790	Sequence 150790,
37	41	82.0	1379	17	US-10-425-114-8021	Sequence 8021, Ap
38	41	82.0	1296	9	US-09-815-242-4090	Sequence 4090, Ap
39	41	82.0	1296	17	US-10-282-122A-7354	Sequence 7354, Ap
40	41	82.0	1318	17	US-10-424-599-9802	Sequence 9802, Ap
41	41	82.0	1318	17	US-10-424-599-12845	Sequence 12845, A
42	41	82.0	1506	14	US-10-014-101-26	Sequence 26, Appl
43	41	82.0	1506	17	US-10-326-184-8	Sequence 8, Appli
44	41	82.0	1506	17	US-10-311-453-26	Sequence 26, Appl
45	41	82.0	1515	14	US-10-014-101-30	Sequence 30, Appl

ALIGNMENTS

RESULT 1  
US-10-087-192-259  
; Sequence 259, Application US/10087192  
; Publication No. US20020182586A1  
; GENERAL INFORMATION:  
; APPLICANT: Morris, David W.  
; APPLICANT: Engelhard, Eric K.  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
; FILE OF INVENTION: CANCER  
; FILE REFERENCE: 529452000122  
; CURRENT APPLICATION NUMBER: US/10/087,192  
; CURRENT FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: US 09/747,377  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/798,586  
; PRIOR FILING DATE: 2001-03-02  
; NUMBER OF SEQ ID NOS: 2059  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 259  
; LENGTH: 65047

```

; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(65047)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-259

```

```

Alignment Scores:
Pred. No.: 2,31e+03 Length: 65047
Score: 45.00 Matches: 8
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 1
Query Match: 90.00% Indels: 0
DB: 13 Gaps: 0

```

```

US-10-014-101B-40 (1-9) x US-10-087-192-259 (1-65047)

```

```

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9

```

```

Db 36758 GTTCTGGAGGTTTGGGGTCTTTTGT 36784

```

#### RESULT 2

```

US-10-029-386-9173
; Sequence 9173, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.

```

```

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 9173
; LENGTH: 503
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```

```

; OTHER INFORMATION: MAP TO AL031277.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: NT HIT: 917705546, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: Q01101, EVALUE 3.00e-05
; OTHER INFORMATION: EST_HUMAN HIT: A1808800.1, EVALUE 1.00e-82
US-10-029-386-9173

```

```

Alignment Scores:
Pred. No.: 44.3 Length: 503
Score: 43.00 Matches: 8
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 1
Query Match: 86.00% Indels: 0
DB: 16 Gaps: 0

```

```

US-10-014-101B-40 (1-9) x US-10-029-386-9173 (1-503)

```

```

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9

```

```

Db 4 GTTCTGGGTGGCGCTGGCCTGTTCTGC 30

```

#### RESULT 3

```

US-10-029-386-22854.
; Sequence 22854, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.

```

```

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 22854
; LENGTH: 789
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL031277.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.94
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.9
; OTHER INFORMATION: SWISSPROT HIT: P57071, EVALUE 3.20e-01
; OTHER INFORMATION: NT HIT: g116168391, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: BG942814.1, EVALUE 0.00e+00
US-10-029-386-22854

```

```

Alignment Scores:
Pred. No.: 69.4 Length: 789
Score: 43.00 Matches: 8
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 1
Query Match: 86.00% Indels: 0
DB: 16 Gaps: 0

```

```

US-10-014-101B-40 (1-9) x US-10-029-386-22854 (1-789)

```

```

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9

```

```

Db 468 GTTCTGGGTGGCGCTGGCCTGTTCTGC 494

```

#### RESULT 4

```

US-10-024-450-3/c
; Sequence 3, Application US/10024450
; Publication No. US20030032606A1
; GENERAL INFORMATION:
; APPLICANT: Huang, Shi
; APPLICANT: Chadwick, Robert B.
; TITLE OF INVENTION: Methods of Detecting and Treating
; FILE REFERENCE: P-LJ 5101
; CURRENT APPLICATION NUMBER: US/10/024,450
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/256,582
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 5868
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (121)...(5277)
US-10-024-450-3

```

```

Alignment Scores:
Pred. No.: 511 Length: 5868
Score: 43.00 Matches: 8
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 1
Query Match: 86.00% Indels: 0
DB: 14 Gaps: 0

```

```

US-10-014-101B-40 (1-9) x US-10-024-450-3 (1-5868)

```

```

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9

```

```

Db 1728 GTTCTGGGTGGCGCTGGCCTGTTCTGC 1702

```

## RESULT 5

US-10-311-455-1795  
; Sequence 1795, Application US/10311455  
; Publication No. US20030143606A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation  
; TITLE OF INVENTION: cytosine methylation  
; FILE REFERENCE: 5013.1014  
; CURRENT APPLICATION NUMBER: US/10/311,455  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: PCT/EP01/07537  
; PRIOR FILING DATE: 2001-07-02  
; PRIOR APPLICATION NUMBER: DE 10032529.7  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 2424  
; SEQ ID NO 1795  
; LENGTH: 6113  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-311-455-1795

Alignment Scores: Length: 6113  
Pred. No.: 533 Matches: 8  
Score: 43.00  
Percent Similarity: 88.89%  
Best Local Similarity: 88.89%  
Query Match: 86.00%  
Indels: 0  
DB: 15 Gaps: 0

US-10-014-101B-40 (1-9) x US-10-311-455-1795 (1-6113)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9

Db 5780 GTTTGGTGGGTAGGTGTTTGT 5806

## RESULT 6

US-10-311-455-654  
; Sequence 654, Application US/10311455  
; Publication No. US20030143606A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation  
; TITLE OF INVENTION: cytosine methylation  
; FILE REFERENCE: 5013.1014  
; CURRENT APPLICATION NUMBER: US/10/311,455  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: PCT/EP01/07537  
; PRIOR FILING DATE: 2001-07-02  
; PRIOR APPLICATION NUMBER: DE 10032529.7  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 2424  
; SEQ ID NO 654  
; LENGTH: 6436  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-311-455-654

Alignment Scores: Length: 6436  
Pred. No.: 561 Matches: 8

Score: 43.00 Matches: 8  
Percent Similarity: 88.89%  
Best Local Similarity: 88.89%  
Query Match: 86.00%  
Indels: 0  
DB: 15 Gaps: 0

US-10-014-101B-40 (1-9) x US-10-311-455-654 (1-6436)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9

Db 5764 GTTTGGTGGGTAGGTGTTTGT 5790

## RESULT 7

US-10-311-455-281  
; Sequence 281, Application US/10311455  
; Publication No. US20030143606A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation  
; TITLE OF INVENTION: cytosine methylation  
; FILE REFERENCE: 5013.1014  
; CURRENT APPLICATION NUMBER: US/10/311,455  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: PCT/EP01/07537  
; PRIOR FILING DATE: 2001-07-02  
; PRIOR APPLICATION NUMBER: DE 10032529.7  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 2424  
; SEQ ID NO 281  
; LENGTH: 7603  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-311-455-281

Alignment Scores: Length: 7603

Pred. No.: 662 Matches: 8  
Score: 43.00  
Percent Similarity: 88.89%  
Best Local Similarity: 88.89%  
Query Match: 86.00%  
Indels: 0  
DB: 15 Gaps: 0

US-10-014-101B-40 (1-9) x US-10-311-455-281 (1-7603)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9

Db 334 GTTTGGTGGGTAGGTGTTTGT 360

## RESULT 8

US-10-240-452-29  
; Sequence 29, Application US/10240452  
; Publication No. US20030162194A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Apoptosis  
; TITLE OF INVENTION: cytosine methylation  
; FILE REFERENCE: 5013.1006  
; CURRENT APPLICATION NUMBER: US/10/240,452  
; CURRENT FILING DATE: 2002-10-02  
; PRIOR APPLICATION NUMBER: PCT/EP01/03969  
; PRIOR FILING DATE: 2001-04-06  
; PRIOR APPLICATION NUMBER: DE 10019058.8  
; PRIOR FILING DATE: 2000-04-06  
; PRIOR APPLICATION NUMBER: DE 10019173.8  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: DE 10032529.7

; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 78  
; SEQ ID NO 29  
; LENGTH: 7603  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-240-452-29

Alignment Scores:  
Pred. No.: 662 Length: 7603  
Score: 43.00 Matches: 8  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 1  
Query Match: 86.00% Indels: 0  
DB: 16 Gaps: 0

US-10-014-101B-40 (1-9) x US-10-240-452-29 (1-7603)

Oy 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
Db 334 GTTTGGTGGGTAGGTGTTTTTGT 360

RESULT 9

US-10-240-453-281  
; Sequence 281, Application US/10240453  
; Publication No. US20030148326A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA  
; TITLE OF INVENTION: Transcription  
; TITLE OF INVENTION: by Means of Assessing the Methylation Status of Genes Associated  
; TITLE OF INVENTION: With DNA Transcription  
; FILE REFERENCE: 5013.1009

US-10-240-453-281  
; CURRENT APPLICATION NUMBER: US/10/240,453  
; CURRENT FILING DATE: 2002-10-02  
; PRIOR FILING DATE: 2002-10-02  
; PRIOR APPLICATION NUMBER: PCT/EP01/03973  
; PRIOR FILING DATE: 2001-04-06  
; PRIOR APPLICATION NUMBER: DE 10019058.8  
; PRIOR FILING DATE: 2000-04-06  
; PRIOR APPLICATION NUMBER: DE 10019173.8  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: DE 10032529.7  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 350  
; SEQ ID NO 281  
; LENGTH: 9504  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-240-453-281

Alignment Scores:  
Pred. No.: 826 Length: 9504  
Score: 43.00 Matches: 8  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 1  
Query Match: 86.00% Indels: 0  
DB: 15 Gaps: 0

US-10-014-101B-40 (1-9) x US-10-240-453-281 (1-9504)

Oy 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
Db 6626 GTTTGGTGGGTAGGTGTTTTTGT 6652

RESULT 10

US-10-087-192-223  
; Sequence 223, Application US/10087192  
; Publication No. US20020182586A1  
; GENERAL INFORMATION:  
; APPLICANT: Morris, David W.  
; APPLICANT: Engelhard, Eric K.  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
; TITLE OF INVENTION: CANCER  
; FILE REFERENCE: 529452000122  
; CURRENT APPLICATION NUMBER: US/10/087,192  
; CURRENT FILING DATE: 2002-03-01  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/747,377  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/798,586  
; NUMBER OF SEQ ID NOS: 2059  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 223  
; LENGTH: 254087  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1)...(254087)  
; OTHER INFORMATION: n = A,T,C or G

US-10-087-192-223

Alignment Scores:  
Pred. No.: 2.18e+04 Length: 254087  
Score: 43.00 Matches: 8  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 1  
Query Match: 86.00% Indels: 0  
DB: 13 Gaps: 0

US-10-014-101B-40 (1-9) x US-10-087-192-223 (1-254087)

Oy 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
Db 241873 GTACIGGGAGGCTGGGTTTGTTC 241899

RESULT 11

US-10-425-115-102522  
; Sequence 102522, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 102522  
; LENGTH: 556  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_25007C.1

US-10-425-115-102522

Alignment Scores:  
Pred. No.: 76.3 Length: 556  
Score: 42.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 87.50% Mismatches: 0  
Query Match: 84.00% Indels: 0  
DB: 18 Gaps: 0

US-10-014-101B-40 (1-9) x US-10-425-115-102522 (1-556)

Qy 2 LeuGlyGlyLeuGlyGlnPheCys 9

Db 510 TTGGGGGGCTGGGGCGTTTCTGT 533

RESULT 12

US-10-027-632-280541/c

; Sequence 280541, Application US/10027632

; Publication No. US20020198371A1

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; FILE OF INVENTION: Polymorphisms in the Human Genome

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632

; PRIOR FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/198,676

; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/193,483

; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23

; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002

; PRIOR FILING DATE: 1999-08-09

; NUMBER OF SEQ ID NOS: 325720

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 280541

; LENGTH: 564

; TYPE: DNA

; ORGANISM: Human

US-10-027-632-280541

Alignment Scores:  
Pred. No.: 77.4 Length: 564  
Score: 42.00 Matches: 8  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 1  
Query Match: 84.00% Indels: 0  
DB: 13 Gaps: 0

US-10-014-101B-40 (1-9) x US-10-027-632-280541 (1-564)

Qy 1 ValLeuGlyGlyLeuGlyGlnPheCys 9

Db 110 GTGCTGGGACCCCTGGGGCAGTTCTGT 84

RESULT 13

US-10-027-632-280541/c

; Sequence 280541, Application US/10027632

; Publication No. US20030204075A9

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; FILE OF INVENTION: Polymorphisms in the Human Genome

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632

; PRIOR FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/198,676

; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/193,483

; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 280541  
; LENGTH: 564  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-280541

Alignment Scores:  
Pred. No.: 77.4 Length: 564  
Score: 42.00 Matches: 8  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 1  
Query Match: 84.00% Indels: 0  
DB: 17 Gaps: 0

US-10-014-101B-40 (1-9) x US-10-027-632-280541 (1-564)

Qy 1 ValLeuGlyGlyLeuGlyGlnPheCys 9

Db 110 GTGCTGGGACCCCTGGGGCAGTTCTGT 84

RESULT 14

US-10-424-599-70249/c

; Sequence 70249, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J

; APPLICANT: Kovalic, David K

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 70249

; LENGTH: 1230

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_34446C.1

US-10-424-599-70249

Alignment Scores:  
Pred. No.: 168 Length: 1230  
Score: 42.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 17.00% Indels: 0  
DB: 17 Gaps: 0

US-10-014-101B-40 (1-9) x US-10-424-599-70249 (1-1230)

Qy 3 GlyGlyLeuGlyGlnPheCys 9

Db 220 GGAGGGTTGGGGCAGTTCTGC 200

RESULT 15

US-10-437-963-82245/c

; Sequence 82245, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

```

; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 82245
; LENGTH: 2139
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_81694C.1
US-10-437-963-82245

Alignment Scores:
Pred. No.: 292 Length: 2139
Score: 42.00 Matches: 6
Percent Similarity: 100.00% Conservative: 3
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 84.00% Indels: 0
DB: 18 Gaps: 0

US-10-014-101B-40 (1-9) x US-10-437-963-82245 (1-2139)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9
Db 2137 ATCGTAGGAGCTCGAGAGGTTTGT 2111

RESULT 16
US-10-221-714A-396
; Sequence 396, Application US/10221714A
; Publication No. US20040048254A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; TITLE OF INVENTION: tumor suppressor genes and oncogenes
; FILE REFERENCE: 5013.1005
; CURRENT APPLICATION NUMBER: US/10/221,714A
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: PCT/EP01/02955
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: DE 10013847.0
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 540
; SEQ ID NO 396
; LENGTH: 7329
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-714A-143

Alignment Scores:
Pred. No.: 994 Length: 7329
Score: 42.00 Matches: 6
Percent Similarity: 100.00% Conservative: 3
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 84.00% Indels: 0
DB: 18 Gaps: 0

US-10-014-101B-40 (1-9) x US-10-437-963-82245 (1-2139)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9
Db 2137 ATCGTAGGAGCTCGAGAGGTTTGT 2111

RESULT 17
US-10-221-714A-143
; Sequence 143, Application US/10221714A
; Publication No. US20040048254A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; TITLE OF INVENTION: tumor suppressor genes and oncogenes
; FILE REFERENCE: 5013.1005
; CURRENT APPLICATION NUMBER: US/10/221,714A
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: PCT/EP01/02955
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: DE 10013847.0
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 540
; SEQ ID NO 143
; LENGTH: 13321
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-714A-143

Alignment Scores:
Pred. No.: 13321 Length: 13321
Score: 42.00 Matches: 6
Percent Similarity: 100.00% Conservative: 3
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 84.00% Indels: 0
DB: 17 Gaps: 0

US-10-014-101B-40 (1-9) x US-10-221-714A-143 (1-13321)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9
Db 10259 GTTATAGTGTCTCGAGAGGTTTGT 10285

RESULT 18
US-10-312-841-1
; Sequence 1, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
; LENGTH: 3673778

```

```

Score: 42.00 Matches: 7
Percent Similarity: 88.89% Conservative: 1
Best Local Similarity: 77.78% Mismatches: 1
Query Match: 84.00% Indels: 0
DB: 17 Gaps: 0

US-10-014-101B-40 (1-9) x US-10-221-714A-396 (1-7329)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9
Db 4553 ATTTTGGCGGTTAGGTGTTTGT 4579

RESULT 17
US-10-221-714A-143
; Sequence 143, Application US/10221714A
; Publication No. US20040048254A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; TITLE OF INVENTION: tumor suppressor genes and oncogenes
; FILE REFERENCE: 5013.1005
; CURRENT APPLICATION NUMBER: US/10/221,714A
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: PCT/EP01/02955
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: DE 10013847.0
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 540
; SEQ ID NO 143
; LENGTH: 13321
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-714A-143

Alignment Scores:
Pred. No.: 13321 Length: 13321
Score: 42.00 Matches: 6
Percent Similarity: 100.00% Conservative: 3
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 84.00% Indels: 0
DB: 17 Gaps: 0

US-10-014-101B-40 (1-9) x US-10-221-714A-143 (1-13321)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9
Db 10259 GTTATAGTGTCTCGAGAGGTTTGT 10285

RESULT 18
US-10-312-841-1
; Sequence 1, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
; LENGTH: 3673778

```

; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Chemically treated genomic DNA (Homo sapiens)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (3294164)  
US-10-312-841-1

Alignment Scores:  
Pred. No.: 4.74e+05 3673778  
Score: 42.00 7  
Percent Similarity: 88.89%  
Best Local Similarity: 77.78%  
Query Match: 84.00%  
DB: 16

US-10-014-101B-40 (1-9) x US-10-312-841-1 (1-3673778)

Qy 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
Db 223774 ATTTGGCGGGTTAGGTGTTTTTGT 223800

RESULT 19

US-10-741-601-10567  
; Sequence 10567, Application US/10741601  
; Publication No. US20040166519A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; FILE REFERENCE: CL001500  
; CURRENT APPLICATION NUMBER: US/10/741.601  
; CURRENT FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 26415  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10567  
; LENGTH: 201  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-741-601-10567

Alignment Scores:  
Pred. No.: 43.2 201  
Score: 41.00 6  
Percent Similarity: 88.89%  
Best Local Similarity: 66.67%  
Query Match: 82.00%  
DB: 18

US-10-014-101B-40 (1-9) x US-10-741-601-10567 (1-201)

Qy 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
Db 6 GTTATGGAGGCATTGGAGCTTTTTCG 32

RESULT 20

US-10-741-600-26062  
; Sequence 26062, Application US/10741600  
; Publication No. US20050026169A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; FILE REFERENCE: CL001499  
; CURRENT APPLICATION NUMBER: US/10/741.600  
; CURRENT FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 73997  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 26062  
; LENGTH: 201  
; TYPE: DNA  
; ORGANISM: Homo sapiens

US-10-741-600-26062

Alignment Scores:  
Pred. No.: 43.2 201  
Score: 41.00 6  
Percent Similarity: 88.89%  
Best Local Similarity: 66.67%  
Query Match: 82.00%  
DB: 19

US-10-014-101B-40 (1-9) x US-10-741-600-26062 (1-201)

Qy 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
Db 6 GTTATGGAGGCATTGGAGCTTTTTCG 32

RESULT 21

US-09-930-213-609/c  
; Sequence 609, Application US/09930213  
; Publication No. US20030170625A1  
; GENERAL INFORMATION:  
; APPLICANT: ROSENTHAL, ANDRE  
; APPLICANT: HINZMANN, BERND  
; APPLICANT: SCHAPER, REINHARD  
; APPLICANT: ZUBER, JOHANNES  
; APPLICANT: TCHE-NITSE, OLEG  
; APPLICANT: GRIPS, MARTIN  
; APPLICANT: HELNEGEL, MARTIN  
; APPLICANT: SCHMITZ, ANNE-CHANTAL  
; APPLICANT: SERS, CHRISTINE  
; TITLE OF INVENTION: DETECTION OF DIFFERENTIAL GENE EXPRESSIONS  
; FILE REFERENCE: ALBRE-14  
; CURRENT APPLICATION NUMBER: US/09/930.213  
; CURRENT FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: DE 10004102.7  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 885  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 609  
; LENGTH: 255  
; TYPE: DNA  
; ORGANISM: Rattus sp.  
US-09-930-213-609

Alignment Scores:  
Pred. No.: 54.7 255  
Score: 41.00 7  
Percent Similarity: 87.50%  
Best Local Similarity: 87.50%  
Query Match: 82.00%  
DB: 10

US-10-014-101B-40 (1-9) x US-09-930-213-609 (1-255)

Qy 2 LeuGlyGlyLeuGlyGlnPheCys 9  
Db 179 CTGGTGGGTTGGGCCCATTTCTGT 156

RESULT 22

US-10-425-115-74928  
; Sequence 74928, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425.115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326

SEQ ID NO 74928  
 LENGTH: 369  
 TYPE: DNA  
 ORGANISM: Zea mays  
 FEATURE:  
 OTHER INFORMATION: Clone ID: MRT4577\_168330C.1  
 US-10-425-115-74928

Alignment Scores:  
 Pred. No.: 79 Length: 369  
 Score: 41.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 82.00% Indels: 0  
 DB: 18 Gaps: 0

US-10-014-101B-40 (1-9) x US-10-425-115-74928 (1-369)

Qy 1 ValLeuGlyLeuGlyGlnPhe 8  
 Db 77 GTGCTGGAGGCTGGGGCAGTTC 100

## RESULT 23

US-10-282-122A-27902/c  
 Sequence 27902, Application US/10282122A  
 Publication No. US20040029129A1

## GENERAL INFORMATION:

APPLICANT: Wang, Liangsu  
 APPLICANT: Zamudio, Carlos  
 APPLICANT: Malone, Cheryl  
 APPLICANT: Haselbeck, Robert  
 APPLICANT: Ohlsen, Kari  
 APPLICANT: Zyskind, Judith  
 APPLICANT: Wall, Daniel  
 APPLICANT: Trawick, John  
 APPLICANT: Carr, Grant  
 APPLICANT: Yamamoto, Robert  
 APPLICANT: Forsyth, R.  
 APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A US/10/282,122A

CURRENT APPLICATION NUMBER: 2003-02-20

PRIOR FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614

SOFTWARE: PatentIn version 3.1

SEQ ID NO 27902

LENGTH: 537

TYPE: DNA

ORGANISM: Mycoplasma pneumoniae

US-10-282-122A-27902

## Alignment Scores:

Pred. No.: 121 Length: 566

Score: 41.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Pred. No.: 115 Length: 537  
 Score: 41.00 Matches: 7  
 Percent Similarity: 88.89% Conservative: 1  
 Best Local Similarity: 77.78% Mismatches: 1  
 Query Match: 82.00% Indels: 0  
 DB: 17 Gaps: 0

US-10-014-101B-40 (1-9) x US-10-282-122A-27902 (1-537)

Qy 1 ValLeuGlyLeuGlyGlnPheCys 9  
 Db 76 GTTGAAGCGGCTGGTCAGTTTTC 50

## RESULT 24

US-10-326-184-15

Sequence 15, Application US/10326184

Publication No. US20030163847A1

GENERAL INFORMATION:

APPLICANT: Monsanto Company

TITLE OF INVENTION: REVERSIBLE MALE STERILITY IN TRANSGENIC PLANTS BY EXPRESSION OF C

FILE REFERENCE: OXIDASE 1

CURRENT APPLICATION NUMBER: US/10/326,184

CURRENT FILING DATE: 2002-12-20

PRIOR APPLICATION NUMBER: US 60/343,129

PRIOR FILING DATE: 2001-12-20

NUMBER OF SEQ ID NOS: 24

SOFTWARE: PatentIn version 3.1

SEQ ID NO 15

LENGTH: 552

TYPE: DNA

ORGANISM: Glycine max

US-10-326-184-15

## Alignment Scores:

Pred. No.: 118 Length: 552

Score: 41.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 82.00% Indels: 0

DB: 16 Gaps: 0

US-10-014-101B-40 (1-9) x US-10-326-184-15 (1-552)

Qy 1 ValLeuGlyLeuGlyGlnPhe 8

Db 12 GTTCTGGAGGCTTAGGGCAATTT 35

## RESULT 25

US-10-326-184-14

Sequence 14, Application US/10326184

Publication No. US20030163847A1

GENERAL INFORMATION:

APPLICANT: Monsanto Company

TITLE OF INVENTION: REVERSIBLE MALE STERILITY IN TRANSGENIC PLANTS BY EXPRESSION OF C

FILE REFERENCE: OXIDASE 1

CURRENT APPLICATION NUMBER: US/10/326,184

CURRENT FILING DATE: 2002-12-20

PRIOR APPLICATION NUMBER: US 60/343,129

PRIOR FILING DATE: 2001-12-20

NUMBER OF SEQ ID NOS: 24

SOFTWARE: PatentIn version 3.1

SEQ ID NO 14

LENGTH: 566

TYPE: DNA

ORGANISM: Glycine max

US-10-326-184-14

## Alignment Scores:

Pred. No.: 121 Length: 566

Score: 41.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 82.00% Indels: 0  
DB: 16 Gaps: 0

US-10-014-101B-40 (1-9) x US-10-326-184-14 (1-566)

QY 1 ValLeuGlyGlyLeuGlyGlnPhe 8  
|||||  
DB 283 GTCTTGAGGCTTAGGGCAATTT 306

RESULT 26  
US-10-326-184-18  
; Sequence 18, Application US/10326184  
; Publication No. US20030163847A1  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Company  
; TITLE OF INVENTION: REVERSIBLE MALE STERILITY IN TRANSGENIC PLANTS BY EXPRESSION OF C  
; FILE REFERENCE: MTC6781.1  
; CURRENT APPLICATION NUMBER: US/10/326,184  
; CURRENT FILING DATE: 2002-12-20  
; PRIOR APPLICATION NUMBER: US 60/343,129  
; PRIOR FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 18  
; LENGTH: 582  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-10-326-184-18

Alignment Scores:  
Pred. No.: 124 Length: 582  
Score: 41.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 82.00% Indels: 0  
DB: 16 Gaps: 0

US-10-014-101B-40 (1-9) x US-10-326-184-18 (1-582)

QY 1 ValLeuGlyGlyLeuGlyGlnPhe 8  
|||||  
DB 58 GTCTTGAGGCTTAGGGCAATTT 81

RESULT 27  
US-10-425-115-94827  
; Sequence 94827, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 94827  
; LENGTH: 588  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_186474C.1  
US-10-425-115-94827

Alignment Scores:  
Pred. No.: 126 Length: 588  
Score: 41.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 82.00% Indels: 0  
DB: 18 Gaps: 0

US-10-014-101B-40 (1-9) x US-10-425-115-94827 (1-598)

QY 1 ValLeuGlyGlyLeuGlyGlnPhe 8  
|||||  
DB 71 GTCTTGAGGCTTAGGGCAATTT 94

RESULT 28  
US-10-326-184-16  
; Sequence 16, Application US/10326184  
; Publication No. US20030163847A1  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Company  
; TITLE OF INVENTION: REVERSIBLE MALE STERILITY IN TRANSGENIC PLANTS BY EXPRESSION OF C  
; FILE REFERENCE: MTC6781.1  
; CURRENT APPLICATION NUMBER: US/10/326,184  
; CURRENT FILING DATE: 2002-12-20  
; PRIOR APPLICATION NUMBER: US 60/343,129  
; PRIOR FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 16  
; LENGTH: 635  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-10-326-184-16

Alignment Scores:  
Pred. No.: 136 Length: 635  
Score: 41.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 82.00% Indels: 0  
DB: 16 Gaps: 0

US-10-014-101B-40 (1-9) x US-10-326-184-16 (1-635)

QY 1 ValLeuGlyGlyLeuGlyGlnPhe 8  
|||||  
DB 108 GTCTTGAGGCTTAGGGCAATTT 131

RESULT 29  
US-10-106-698-1224  
; Sequence 1224, Application US/10106698  
; Publication No. US20030109690A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide  
; FILE REFERENCE: PA005P1  
; CURRENT APPLICATION NUMBER: US/10/106,698  
; CURRENT FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: PCT/US00/26524  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US 60/157,137  
; PRIOR FILING DATE: 1999-09-29  
; PRIOR APPLICATION NUMBER: US 60/163,280  
; PRIOR FILING DATE: 1999-11-03  
; NUMBER OF SEQ ID NOS: 8564  
; SOFTWARE: PatentIn Ver. 3.0  
; SEQ ID NO 1224  
; LENGTH: 1088  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (4)..(4)  
; OTHER INFORMATION: n equals a.t.g. or c  
; NAME/KEY: misc\_feature  
; LOCATION: (158)..(158)  
; OTHER INFORMATION: n equals a.t.g. or c

US-10-106-698-1224  
Alignment Scores:  
Pred. No.: 232 Length: 1088  
Score: 41.00 Matches: 7  
Percent Similarity: 87.50% Conservative: 0  
Best Local Similarity: 87.50% Mismatches: 1  
Query Match: 82.00% Indels: 0  
DB: 15 Gaps: 0

US-10-014-101B-40 (1-9) x US-10-106-698-1224 (1-1088)

Qy 2 ValLeuGlyLeuGlyGlnPheCys 9  
Db 59 CTCGGTGGCGCTCGCTCGTTCGTC 82

## RESULT 30

US-10-425-114-10886  
; Sequence 10886, Application US/10425114  
; Publication No. US2004003488A1

## GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 10886

; LENGTH: 1141

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: 700985454\_FLI

US-10-425-114-10886

Alignment Scores:  
Pred. No.: 243 Length: 1141  
Score: 41.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 82.00% Indels: 0  
DB: 17 Gaps: 0

US-10-014-101B-40 (1-9) x US-10-425-114-10886 (1-1141)

Qy 1 ValLeuGlyLeuGlyGlnPhe 8  
Db 58 GTTCTAGGGGGCTTGCCCAATT 81

## RESULT 31

US-10-424-599-1725

; Sequence 1725, Application US/10424599

; Publication No. US20040031072A1

## GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic, David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424, 599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 1725

; LENGTH: 1142

; TYPE: DNA

; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_101557C.1  
US-10-424-599-1725

Alignment Scores:  
Pred. No.: 243 Length: 1142  
Score: 41.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 82.00% Indels: 0  
DB: 17 Gaps: 0

US-10-014-101B-40 (1-9) x US-10-424-599-1725 (1-1142)

Qy 1 ValLeuGlyLeuGlyGlnPhe 8

Db 59 GTTCTAGGGGGCTTGCCCAATT 82

## RESULT 32

US-10-363-345A-20791

; Sequence 20791, Application US/10363345A

; Publication No. US20040234960A1

## GENERAL INFORMATION:

; APPLICANT: Alexander Olek  
; APPLICANT: Christian Piepenbrock

; APPLICANT: Kurt Berlin

; TITLE OF INVENTION: Method for determining the degree of methylation of defined

; FILE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3

; FILE REFERENCE: E01/1227

; CURRENT APPLICATION NUMBER: US/10/363,345A

; CURRENT FILING DATE: 2003-03-03

; NUMBER OF SEQ ID NOS: 40712

; SEQ ID NO 20791

; LENGTH: 1234

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

; OTHER INFORMATION: CpG-island No: 20791

US-10-363-345A-20791

Alignment Scores:  
Pred. No.: 263 Length: 1234  
Score: 41.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 3  
Best Local Similarity: 66.67% Mismatches: 0  
Query Match: 82.00% Indels: 0  
DB: 18 Gaps: 0

US-10-014-101B-40 (1-9) x US-10-363-345A-20791 (1-1234)

Qy 1 ValLeuGlyLeuGlyGlnPheCys 9  
Db 532 GTCGTCGGCGGCTTGAGTTTGC 558

## RESULT 33

US-10-363-345A-20792/c

; Sequence 20792, Application US/10363345A

; Publication No. US20040234960A1

## GENERAL INFORMATION:

; APPLICANT: Alexander Olek  
; APPLICANT: Kurt Berlin

; APPLICANT: Christian Piepenbrock

; TITLE OF INVENTION: Method for determining the degree of methylation of defined

; FILE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3

; FILE REFERENCE: E01/1227

; CURRENT APPLICATION NUMBER: US/10/363,345A

; CURRENT FILING DATE: 2003-03-03

; NUMBER OF SEQ ID NOS: 40712

; SEQ ID NO 20792

; LENGTH: 1234

; TYPE: DNA



; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 8021  
; LENGTH: 1279  
; TYPE: DNA  
; ORGANISM: Glycine max

; FEATURE:  
; OTHER INFORMATION: Clone ID: 700731265\_FLI  
US-10-425-114-8021

Alignment Scores: 272 Length: 1279  
Pred. No.: 41.00 Matches: 8  
Score: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 82.00% Indels: 0  
DB: 17 Gaps: 0

US-10-014-101B-40 (1-9) x US-10-425-114-8021 (1-1279)

QY 1 ValLeuGlyGlyLeuGlyGlnPhe 8  
|||||  
DB 21 GTACTGGAGGCTAGTCACTTT 44  
|||||

RESULT 38  
US-09-815-242-4090  
; Sequence 4090, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; TITLE OF INVENTION: Prokaryotes  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4090  
; LENGTH: 1296  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa

US-09-815-242-4090  
Alignment Scores: 276 Length: 1296  
Pred. No.: 41.00 Matches: 7  
Score: 100.00% Conservative: 1  
Best Local Similarity: 100.00% Mismatches: 1  
Query Match: 82.00% Indels: 0  
DB: 17 Gaps: 0

US-10-014-101B-40 (1-9) x US-10-425-114-8021 (1-1279)

QY 1 ValLeuGlyGlyLeuGlyGlnPhe 8  
|||||  
DB 21 GTACTGGAGGCTAGTCACTTT 44  
|||||

Percent Similarity: 88.89% Conservative: 1  
Best Local Similarity: 77.78% Mismatches: 1  
Query Match: 92.00% Indels: 0  
DB: 9 Gaps: 0

US-10-014-101B-40 (1-9) x US-09-815-242-4090 (1-1296)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
|||||  
DB 1081 GTGATCGGCGCTCGCGGCTTCTGC 1107  
|||||

RESULT 39  
US-10-282-122A-7354  
; Sequence 7354, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 7354  
; LENGTH: 1296  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa

US-10-282-122A-7354  
Alignment Scores: 276 Length: 1296  
Pred. No.: 41.00 Matches: 7  
Score: 88.89% Conservative: 1  
Best Local Similarity: 77.78% Mismatches: 1  
Query Match: 82.00% Indels: 0  
DB: 17 Gaps: 0

US-10-014-101B-40 (1-9) x US-10-282-122A-7354 (1-1296)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
|||||

Db 1081 GTGATCGCGCGCTCGCGGCTTCTGC 1107

RESULT 40

US-10-424-599-9802

; Sequence 9802, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 9802

; LENGTH: 1318

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_108862C.1

US-10-424-599-9802

Alignment Scores:

Pred. No.:	281	Length:	1318
Score:	41.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	82.00%	Indels:	0
DB:	17	Gaps:	0

US-10-014-101b-40 (1-9) x US-10-424-599-9802 (1-1318)

Qy 1 ValLeuGlyLeuGlyGlnPhe 8

Db 778 GTTCTTGAGGCTTAGGGCAATTT 801

Search completed: February 18, 2005, 10:15:49

Job time : 633.26 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 18, 2005, 02:20:24 ; Search time 1061.28 Seconds  
(without alignments)  
322.797 Million cell updates/sec

Title: US-10-014-101b-40  
Perfect score: 50  
Sequence: 1 VLGLGQFC 9

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
	Ygapop 10.0	Ygapext 0.5
	Fgapop 6.0	Fgapext 7.0
	Delop 6.0	Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-O=cgn2 1/USPTO.apool/US10014101/runat 16022005 075846 7952/app query.fasta\_1.796  
-DB-EST -QPMT-fastap -SUFFIX=rat -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10014101 @CGN 1 1 5453 @runat 16022005 075846 7952 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: \*  
1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_hic.\*  
4: gb\_est3.\*  
5: gb\_est4.\*  
6: gb\_est5.\*  
7: gb\_est6.\*  
8: gb\_gsl1.\*  
9: gb\_gsl2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	50	100.0	911	8 BZ178637	BZ178637 CH230-456
C 2	47	94.0	562	7 CK375056	CK375056 lai44c10.
C 3	46	92.0	415	9 CU895276	CU895276 abg31h01.
4	46	92.0	951	5 BU169208	BU169208 AGENCOURT
5	45	90.0	206	7 CF558245	CF558245 1115035C0
C 6	45	90.0	494	2 BE944702	BE944702 UI-M-BH3
C 7	45	90.0	501	2 BE949131	BE949131 UI-M-BH3
C 8	45	90.0	590	5 BQ638095	BQ638095 hdi8d07.y
C 9	44	88.0	528	8 AQ807467	AQ807467 HS_3207_B

C 10	44	88.0	528	8	AQ807479	AQ807479 HS_3207_B
11	44	88.0	699	6	BY748808	BY748808 BY748808
C 12	44	88.0	846	4	BI916220	BI916220 603178115
C 13	44	88.0	1007	5	BQ439107	BQ439107 AGENCOURT
C 14	43	86.0	276	2	BB608644	BB608644 BB608644
C 15	43	86.0	370	4	BG942814	BG942814 ax30a01.x
C 16	43	86.0	425	1	AI552042	AI552042 mm73g06.y
C 17	43	86.0	433	1	AA072623	AA072623 mm73g06.y
C 18	43	86.0	443	9	CE667699	CE667699 tigr-g88-
C 19	43	86.0	498	2	AW180619	AW180619 MGA0762f
C 20	43	86.0	516	6	CD217252	CD217252 TgESTzy12
C 21	43	86.0	517	5	BU699322	BU699322 Li2in1125
C 22	43	86.0	530	4	BI848950	BI848950 471549 MA
C 23	43	86.0	576	5	EX643028	EX643028 DXFZP781G
C 24	43	86.0	594	5	BP768145	BP768145 BP768145
C 25	43	86.0	650	2	BB618995	BB618995 BB618995
C 26	43	86.0	652	8	BZ952846	BZ952846 CH240.125
C 27	43	86.0	676	8	BH099003	BH099003 RPCI-24-3
C 28	43	86.0	714	2	BE176609	BE176609 RC3-HT058
C 29	43	86.0	827	9	CL714569	CL714569 OR_BBA003
C 30	43	86.0	1143	5	BQ050317	BQ050317 AGENCOURT
C 31	42	84.0	245	2	BB240512	BB240512 BB240512
C 32	42	84.0	253	2	BB603301	BB603301 BB603301
C 33	42	84.0	400	1	AL361639	AL361639 AL361639
C 34	42	84.0	413	8	AZ141369	AZ141369 SP_0042_A
C 35	42	84.0	429	8	AQ702310	AQ702310 HS_5332_B
C 36	42	84.0	445	5	BY451956	BY451956 BY451956
C 37	42	84.0	446	4	BM271588	BM271588 TgESTzyB4
C 38	42	84.0	446	5	BW523847	BW523847 BW523847
C 39	42	84.0	452	7	CN605793	CN605793 PEI0290_P
C 40	42	84.0	476	7	CO072048	CO072048 GR_Ea31A
C 41	42	84.0	495	7	CO664303	CN664303 A0811d10-
C 42	42	84.0	500	7	CN605683	CN605683 PEI0091_P
C 43	42	84.0	530	8	AQ189600	AQ189600 HS_3219_B
C 44	42	84.0	537	8	BH330237	BH330237 CH230-58M
C 45	42	84.0	555	4	BG883465	BG883465 fp29g12.x

#### ALIGNMENTS

RESULT 1

BZ178637

LOCUS

DEFINITION

BZ178637 911 bp DNA linear GSS 11-OCT-2002

CH230-456M15.TV CHORI-230 Segment 2 Rattus norvegicus genomic clone

CH230-456M15, genomic survey sequence.

BZ178637 GI:23822768

VERSION

KEYWORDS

SOURCE

ORGANISM

Rattus norvegicus (Norway rat)

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 911)

Zhao,S., Shetty,J., Shatsman,S., Teegaye,G., Geer,K.,

Sivatsbeyn,A., Gebregorgis,E., Overton,L., Russell,D., Chen,D.,

Riggs,F., de Jong,P. and Fraser,C.M.

Rat BAC End Sequences from Library CHORI-230 MboI segment

Unpublished (1999)

Other GSSs: CH230-456M15.TJ

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the rat BAC library CHORI-230

(<http://www.chori.org/bacpac/rat230.htm>). For BAC library

availability, please contact Pieter de Jong (pdejong@mail.cho.org).

Clones may be purchased from BACPAC Resources

(<http://www.chori.org/bacpac/orering/information.htm>). BAC end

page: [http://www.tigr.org/tdb/bac\\_ends/rat/bac\\_end\\_intro.html](http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html)

Plate: 456 row: M column: 15  
Seq primer: T7  
Class: BAC ends.

FEATURES  
source  
1. .911  
Location/Qualifiers  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/strain="BN/SsNHsd/MCW"  
/db\_xref="taxon:10116"  
/clone="CH230-456M15"  
/sex="Female"  
/cell\_type="Brain"  
/clone\_lib="CHORI-230 Segment 2"  
/notes="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;  
CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by  
Pieter de Jong"

## ORIGIN

Alignment Scores:  
Pred. No.: 75.1 Length: 562  
Score: 50.00 Matches: 9  
Percent Similarity: 100.00% Conservativeness: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0

US-10-014-101B-40 (1-9) x BZ178637 (1-911)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
Db 17 GTTCTGGGGGACTAGTCAATTTGT 43

## RESULT 2

CK375056/c  
LOCUS CK375056 562 bp mRNA linear EST 23-DEC-2003  
DEFINITION lai44c10.y1 Gastric Epithelial Progenitor 2 Mus musculus cDNA 5', similar to SW: TLM\_MOUSE Pf17408 TLM PROTEIN ; mRNA sequence.

ACCESSION CK375056  
VERSION CK375056.1 GI:40340298  
KEYWORDS EST.

SOURCE Mus musculus (house mouse)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 562)  
Tidwell, R., Clifton, S., Marra, M., Hillier, L., Pape, D., Martin, J., Wylie, T., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Ronko, I., Tsagarelis, V., Belaygorod, L., Grow, A., Maguire, L., Waterston, R., and Wilson, R.

WashU Stem Cell EST Project

Unpublished (2002)

Contact: Jeff Gordon and Mike Lovett

WashU, Human Genetics Division

Washington University School of Medicine

1st strand of cDNA was synthesized with reverse transcriptase and oligo(dT) beads, then cDNA was amplified by PCR using modified SMART primers. The final cDNA was cloned in pAMP1 vector in an annealing reaction with Uracil DNA Glycosylase (UDG). Library constructed by Y. Korshunova and M. Lovett. Library materials provided by Mills JC & Gordon JI.

Possible reversed clone: similarity on wrong strand

Seq primer: -40RP from Gibco.

## FEATURES

source  
1. .562  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/tissue\_type="Laser-captured isthmus cells from toxi176 transgenic mice"  
/dev stage="adult"  
/lab\_host="DHSalpha"  
/clone\_lib="Gastric Epithelial Progenitor 2"  
/notes="Vector: pAMP1; 1st strand of cDNA was synthesized

## ORIGIN

Alignment Scores:  
Pred. No.: 162 Length: 562  
Score: 47.00 Matches: 8  
Percent Similarity: 100.00% Conservativeness: 1  
Best Local Similarity: 88.89% Mismatches: 0  
Query Match: 94.00% Indels: 0  
DB: 7 Gaps: 0

US-10-014-101B-40 (1-9) x CK375056 (1-562)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
Db 448 CTCCTTAGGGGCTTGGTCAATTTGC 422

## RESULT 3

CL895276/c

LOCUS CL895276 415 bp DNA linear GSS 30-AUG-2004  
DEFINITION abg31h01.y1 Soybean random, unfiltered genomic library Glycine max genomic, genomic survey sequence.

ACCESSION CL895276

VERSION CL895276.1 GI:51644968

KEYWORDS GSS.

SOURCE Glycine max (soybean)

## ORGANISM

Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

1 (bases 1 to 415)

Numberg, A., Bedell, J. A., Citek, R. W., Robbins, D., McMenamy, J., Peterson, S., Jones, J., Fries, J., Budiman, M. A., Nguyen, H. and

Stacey, G.

Methylation filtered genomic sequences from Glycine max

## JOURNAL

Unpublished (2004)

Contact: Gary Stacey

University of Missouri

108 Waters Hall, Columbia, MO 65211, USA

Tel: 573-884-1267

Fax: 573-882-0588

Email: stacey@missouri.edu

Lidid: 230

Class: shotgun.

## FEATURES

Location/Qualifiers  
source  
1. .415  
/organism="Glycine max"  
/mol\_type="genomic DNA"  
/cultivar="Williams 82"  
/db\_xref="taxon:3847"  
/tissue\_type="Young leaves"  
/clone\_lib="Soybean random, unfiltered genomic library"  
/note="Vector: pOT2; Site 1: BstXI; Randomly sheared and genomic DNA ranging from 0.7-1.5 kb were end repaired and ligated to BstXI linkers prior to cloning in BstXI-cut pOT2. LibID: 230"

## ORIGIN

Alignment Scores:  
Pred. No.: 179 Length: 415  
Score: 46.00 Matches: 8  
Percent Similarity: 100.00% Conservativeness: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 92.00% Indels: 0  
DB: 9 Gaps: 0

US-10-014-101B-40 (1-9) x CL895276 (1-415)

with reverse transcriptase and oligo(dT) beads, then cDNA was amplified by PCR using modified SMART primers. The final cDNA was cloned in pAMP1 vector in annealing reaction with Uracil DNA Glycosylase (UDG). Library constructed by Y. Korshunova and M. Lovett. Library materials provided by Mills JC & Gordon JI."

```

Qy 2 LeuGlyGlyLeuGlyGlnPheCys 9
Db 267 TTGGGGGGGACTTGCTGCAATTTTC 244

RESULT 4
BU169208
LOCUS BU169208 951 bp mRNA linear EST 04-SEP-2002
DEFINITION AGENCOURT_8099372 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6254236
5', mRNA sequence.
ACCESSION BU169208
VERSION BU169208.1 GI:226683192
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 951)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2403 row: g column: 05
High quality sequence start: 23
High quality sequence stop: 735.
FEATURES
source
1..951
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6254236"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_102"
/notes="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies). Note: this is a NIH_MGC Library."

ORIGIN
Alignment Scores:
Pred. No.: 436 Length: 951
Score: 46.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.00% Indels: 0
DB: 5 Gaps: 0

US-10-014-101B-40 (1-9) x BU169208 (1-951)

Qy 2 LeuGlyGlyLeuGlyGlnPheCys 9
Db 916 TTGGGGGGGCTTGCTGCAATTTTC 939

RESULT 5
CF558245
LOCUS CF558245 206 bp mRNA linear EST 23-SEP-2003
DEFINITION 1115035C04.x2 C. reinhardtii CC-1690 (mt+), Stress III
(normalized), Lambda Zap II Chlamydomonas reinhardtii cDNA 3', mRNA
sequence.
ACCESSION CF558245

```

```

VERSION CF558245.1 GI:34984328
KEYWORDS Chlamydomonas reinhardtii
SOURCE Chlamydomonas reinhardtii
ORGANISM Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
REFERENCE 1 (bases 1 to 206)
AUTHORS Groseman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C.,
Merchant, S., Quinn, J. and Shrager, J.
TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants. Project: 1115
JOURNAL Unpublished (2003)
COMMENT Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu
High quality sequence stop: 206.
FEATURES
source
1..206
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690 (mt+), Stress III
(normalized), Lambda Zap II"
/notes="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
XhoI; Stress condition III library, constructed by
Jeanette Quinn and Chung-wen Chang, combines cDNAs from
CC-1690 cells grown to mid-log phase in copper-free TAP
medium (see Quinn and Merchant (1998) Methods in
Enzymology, 297:263-279) in a shaking (250 rpm) illuminated
(about 100 micromole/m2/sec) incubator at 22 C (see Quinn
and Merchant (1998) Methods in Enzymology, 297:263-279);
CC-1690 cells grown to mid-log phase in low Fe (1
micromolar Fe) TAP medium (see La Fontaine S, Quinn JM,
Nakamoto SS, Page MD, Gohre V, Moseley JL, Kropat J,
Merchant S, LaFontaine et al. (2002) Eukaryotic Cell,
1:736-757) in a shaking illuminated incubator (same
conditions as above) CC-1690 cells were grown to mid-log
phase in TAP medium in a shaking illuminated incubator to
a density of 8x10e6 cells/ml. The flask was transferred to
a shaking platform (200 rpm) at room temperature (23C) 12
micromole/m2/sec illumination and bubbled in a stoppered
flask with 98% nitrogen, 2% CO2 gas mixture for 1 hour
prior to harvesting for RNA isolation (as per Quinn JM,
Barraco P, Eriksson M, Merchant S, Quinn et al. (2000) JBC
275:6080-6089); CC-1690 cells grown to mid-log phase
(3x10e6 cells/ml) in TAP medium in a shaking (150 rpm)
illuminated (70 mole photon/m2/sec) incubator at 27 C.
Cells were diluted to 1x10e6 cells/ml, transferred to high
light (11000 mol photon/m2/s) with shaking (150 rpm) and
sampled at (0.5, 1.2, 4, 6, 12 hrs); CC-1690 cells grown to
mid-log phase in HS medium in a shaking (150 rpm)
illuminated (70 mole photon/m2/sec) incubator at 27 C.
Cells were diluted to 1x10e6 cells/ml, transferred to high
light (11000 mol photon/m2/s) with shaking (150 rpm) and
sampled at (0.5, 1.2, 4, 6, 12 hrs). PolyA mRNA was purified
from each sample, pooled and cDNA synthesized (see Shrager
et al, Plant Physiol. 131, 401-408 for details). The cDNA
was directionally cloned into lambda Zap II (Stratagene)
in the EcoRI (5') and XhoI (3') sites. pBluescript II SK-
plasmids were excised from the lambda Zap clones by
superinfection with Exassist (Stratagene) phage. The
library was normalized using method 4 described in Bonaldo
et al., (1996) Genome Research 6: 791-806."
ORIGIN
Alignment Scores:
Pred. No.: 130 Length: 206

```

Score: 45.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 2  
 Best Local Similarity: 77.78% Mismatches: 0  
 Query Match: 90.00% Indels: 0  
 DB: 7 Gaps: 0

US-10-014-101B-40 (1-9) x CF558245 (1-206)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9

Db 156 ATCTGGGGGATTGGGGCGGTCTGC 182

RESULT 6

BE944702/c

LOCUS

DEFINITION UI-M-BH3-awg-a-04-0-UI.s1 NIH\_BMAP\_M\_S4 Mus musculus cDNA clone

ACCESSION BE944702

VERSION BE944702.1 GI:10522461

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 494)

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene

discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

PUBMED 8889548

COMMENT

Contact: Chin, H

National Institute of Mental Health

6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD

20892-9643, USA

Tel: 301 443 1706

Fax: 301 443 9890

Email: mEST@mail.nih.gov

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the olfactory bulbs tissue cDNA library Preparation: M.B. Soares Lab Clone distribution: Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It should be noted that Bento Soares is

generating a small number of additional specialized non-redundant arrays of BMAP cDNAs whose availability will be considered under appropriate and limited collaborative arrangements The following repetitive elements were found in this cDNA sequence: 265-494,

>MTD#LTR/MaLR

Seq primer: M13 Forward

POLYA=Yes.

FEATURES

source

1. 494

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="UI-M-BH3-awg-a-04-0-UI"

/dev\_stage="27-32 days"

/lab\_host="DH10B (Life Technologies)"

/clone\_lib="NIH\_BMAP\_M\_S4"

/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The NIH\_BMAP\_M\_S4 library is a subtracted library of a series, ultimately derived from a mixture of individually tagged normalized libraries from ten regions of the mouse brain (cerebellum, brain stems, olfactory bulbs, hypothalamus, cortex, amygdala, basal ganglia, pineal gland, striatum, hippocampus) after a series of subtractions to reduce the representation of cDNAs from which ESTs had already been generated. The following serially subtracted libraries

were generated in this process: NIH\_BMAP\_M\_S4, NIH\_BMAP\_M\_S3.3, NIH\_BMAP\_M\_S3.2, NIH\_BMAP\_M\_S3.1, NIH\_BMAP\_M\_S2, NIH\_BMAP\_M\_S1. The subtracted library (NIH\_BMAP\_M\_S4) was constructed as follows: PCR amplified cDNA inserts from NIH\_BMAP\_M\_S3.3, NIH\_BMAP\_M\_S3.2, and NIH\_BMAP\_M\_S3.1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with a pool of the NIH\_BMAP\_M\_S3.3, NIH\_BMAP\_M\_S3.2, and NIH\_BMAP\_M\_S3.1 libraries in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the NIH\_BMAP\_M\_S4 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)

TAG\_TISSUE=olfactory-bulbs

TAG\_LIB=NIH\_BMAP\_M\_S4

TAG\_SEQ=CAAGG"

ORIGIN

Alignment Scores: 332 Length: 494

Pred. No.: 45.00 Matches: 8

Score: 88.89% Conservative: 0

Best Local Similarity: 88.89% Mismatches: 1

Query Match: 90.00% Indels: 0

DB: 2 Gaps: 0

US-10-014-101B-40 (1-9) x BE944702 (1-494)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9

Db 238 GTGTTAGGGGTTTGGGGATTTCGT 212

RESULT 7

BE949131/c

LOCUS

DEFINITION

UI-M-BH3-avh-e-05-0-UI.s1 NIH\_BMAP\_M\_S4 Mus musculus cDNA clone

ACCESSION BE949131

VERSION BE949131.1 GI:10526890

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 501)

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene

discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

PUBMED 8889548

COMMENT

Contact: Chin, H

National Institute of Mental Health

6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD

20892-9643, USA

Tel: 301 443 1706

Fax: 301 443 9890

Email: mEST@mail.nih.gov

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the olfactory bulbs tissue cDNA library Preparation: M.B. Soares Lab Clone distribution: Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It should be noted that Bento Soares is generating a small number of additional specialized non-redundant arrays of BMAP cDNAs whose availability will be considered under appropriate and limited collaborative arrangements The following repetitive elements were found in this cDNA sequence: 265-501,

>MTDLTR/MALR  
Seq primer: M13 Forward  
POLYA=yes.

FEATURES  
source  
1. .501  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UI-M-BH3-avh-e-05-01-UI"  
/dev\_stage="27-32 days"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="NIH\_BMAP\_M\_S4"  
/note="Vector: p773D-Fac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The NIH\_BMAP\_M\_S4 library is a subtracted library of a series, ultimately derived from a mixture of individually tagged normalized libraries from ten regions of the mouse brain (cerebellum, brain stems, olfactory bulbs, hypothalamus, cortex, amygdala, basal ganglia, pineal gland, striatum, hippocampus) after a series of subtractions to reduce the representation of cDNAs from which ESTs had already been generated. The following serially subtracted libraries were generated in this process: NIH\_BMAP\_M\_S4, NIH\_BMAP\_M\_S3.3, NIH\_BMAP\_M\_S3.2, NIH\_BMAP\_M\_S3.1, NIH\_BMAP\_M\_S2, NIH\_BMAP\_M\_S1. The subtracted library (NIH\_BMAP\_M\_S4) was constructed as follows: PCR amplified cDNA inserts from NIH\_BMAP\_M\_S3.3, NIH\_BMAP\_M\_S3.2, and NIH\_BMAP\_M\_S3.1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with a pool of the NIH\_BMAP\_M\_S3.3, NIH\_BMAP\_M\_S3.2, and NIH\_BMAP\_M\_S3.1 libraries in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the NIH\_BMAP\_M\_S4 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)  
TAG TISSUE=olfactory-bulbs  
TAG LIB=NIH\_BMAP\_M\_S4  
TAG\_SEQ=CATGG"

ORIGIN  
Alignment Scores:  
Pred. No.: 337 Length: 501  
Score: 45.00 Matches: 8  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 1  
Query Match: 90.00% Indels: 0  
DB: 2 Gaps: 0

US-10-014-101B-40 (1-9) x BE949131 (1-501)

QY 1 ValLeuGlyLeuGlyGlnPheCys 9  
|||||  
Db 238 GTGTAGGGGTTTGGGGGATTTTGT 212  
|||||

RESULT 8  
BQ638095/c 590 bp mRNA linear EST 15-JUL-2002  
LOCUS hd18d07.y1 Human Retina cDNA (Un-normalized, unamplified): hd/he  
DEFINITION Homo sapiens cDNA clone hd18d07 5', mRNA sequence.  
ACCESSION BQ638095  
VERSION BQ638095.1 GI:21762554.  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 590)  
AUTHORS Wistow,G., Bernstein,S.L., Wyatt,M.K., Ray,S., Behal,A., Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.

Expressed sequence tag analysis of human retina for the NEIBank Project: Retbindin, an abundant, novel retinal cDNA and alternative splicing of other retina-preferred gene transcripts  
Mol. Vis. 8 (4), 196-204 (2002)  
22103461  
12107411  
Contact: Wistow G  
Section on Molecular Structure and Function  
National Eye Institute  
6/331, NIH, Bethesda, MD 20892-2740, USA  
Tel: 301 402 3452  
Fax: 301 496 0078  
Email: graeme@helix.nih.gov  
Plate: 18 row: d column: 07  
Seq primer: M13RPI reverse primer (ABI).  
Location/Qualifiers  
1. .590  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="hd18d07"  
/tissue\_type="Retina"  
/dev\_stage="Adult"  
/lab\_host="EMDH10B"  
/clone\_lib="Human Retina cDNA (Un-normalized, unamplified): hd/he"  
/note="Organ: Eye; Vector: pSPORT1; Neural retina tissue was dissected from two 80 year old donors with no observed eye disease. 100ug of total RNA was used for library construction. A directionally cloned cDNA library in the pSPORT1 vector (Life Technologies) was constructed at Bioserve Biotechnology (Laurel MD) essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's Instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adaptor  
[5'-pGACVAGTTCCTAGATCGGAGCGCGGCC(7)15-3']. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."

ORIGIN  
Alignment Scores:  
Pred. No.: 401 Length: 590  
Score: 45.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 2  
Best Local Similarity: 77.78% Mismatches: 0  
Query Match: 90.00% Indels: 0  
DB: 5 Gaps: 0

US-10-014-101B-40 (1-9) x BQ638095 (1-590)

QY 1 ValLeuGlyLeuGlyGlnPheCys 9  
:::|||||  
Db 115 ATCTTGGGGGCTGGGGATTTTGT 89  
|||||

RESULT 9  
AQ807467/c 528 bp DNA linear GSS 10-AUG-1999  
LOCUS HS\_3207\_B1\_C04\_T7C CIT Approved Human Genomic Sperm Library D Homo  
DEFINITION sapiens genomic clone Plate=3207 Col=7 Row=F, genomic survey sequence.  
ACCESSION AQ807467  
VERSION AQ807467.1 GI:5726709  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 528)  
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.

**TITLE** Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

**JOURNAL** Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

**MEDLINE** 99380589

**PUBMED** 10449764

**COMMENT** Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Clones may be purchased from Research Genetics (info@resgen.com).  
BAC end Web Server: http://www.htsc.washington.edu  
Plate: 3207 row: F column: 7  
Seq primer: T7  
Class: BAC ends  
High quality sequence stop: 528.

**FEATURES**  
source  
1..528  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="plate=3207 Col=7 Row=F"  
/sex="male"  
/clone\_lib="CIT Approved Human Genomic Sperm Library D"  
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in E-Coli DH10B"

**ORIGIN**

Alignment Scores:  
Pred. No.: 547 Length: 528  
Score: 44.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 87.50% Mismatches: 0  
Query Match: 88.00% Indels: 0  
DB: 8 Gaps: 0

US-10-014-101b-40 (1-9) x AQ807467 (1-528)

QY 2 LeuGlyGlyLeuGlyGlnPheCys 9  
:::|||||  
DB 506 ATAGGTGGCTGGGGCAATTCGT 483

RESULT 10  
AQ807479/c  
LOCUS  
DEFINITION  
HS 3207 B1.D04.T7C CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3207 Col=7 Row=H, genomic survey sequence.  
ACCESSION  
AQ807479  
VERSION  
AQ807479.1 GI:5726721  
KEYWORDS  
GSS.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 528)  
AUTHORS  
Mahairas G.G., Wallace J.C., Smith K., Swartzell S., Holzman T., Keller A., Shaker R., Furlong J., Young J., Zhao S., Adams M.D. and Hood L.

**TITLE** Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

**JOURNAL** Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

**MEDLINE** 99380589

**PUBMED** 10449764

**COMMENT** Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu

Clones may be purchased from Research Genetics (info@resgen.com).  
BAC end Web Server: http://www.htsc.washington.edu  
Plate: 3207 row: H column: 7  
Seq primer: T7  
Class: BAC ends  
High quality sequence stop: 528.

**FEATURES**  
source  
1..528  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="plate=3207 Col=7 Row=H"  
/sex="male"  
/clone\_lib="CIT Approved Human Genomic Sperm Library D"  
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in E-Coli DH10B"

**ORIGIN**

Alignment Scores:  
Pred. No.: 547 Length: 528  
Score: 44.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 87.50% Mismatches: 0  
Query Match: 88.00% Indels: 0  
DB: 8 Gaps: 0

US-10-014-101b-40 (1-9) x AQ807479 (1-528)

QY 2 LeuGlyGlyLeuGlyGlnPheCys 9  
:::|||||  
DB 506 ATAGGTGGCTGGGGCAATTCGT 483

RESULT 11  
BY748808  
LOCUS  
DEFINITION  
BY748808 RIKEN full-length enriched, NOD-derived cDNA clone F630207H20 5', mRNA sequence.  
ACCESSION  
BY748808  
VERSION  
BY748808.1 GI:27177838  
KEYWORDS  
EST.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1 (bases 1 to 699)  
AUTHORS  
Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S., Nikaide I., Oato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H., Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T., Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W., Blake J.A., Bradt D., Bruscia V., Chothia C., Corbani L.E., Cousins S., Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S., Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J., Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D., Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L., Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A., Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H., Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G., Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S., Ravasi T., Reed J.C., Reid J.J., Reid J., Ring B.Z., Ringwald M., Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K., Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M., Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C., Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L., Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N., Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K., Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S., Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I., Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A., Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J., Birney E. and Hayaehizaki Y.

**TITLE**  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs

**JOURNAL**  
Nature 420, 563-573 (2002)

**MEDLINE**  
22354683

**PUBMED**  
12466851

**COMMENT**  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9226  
Fax: 81-45-503-9212  
Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/  
Adachi, J., Aizawa, K., Akumura, T., Arakawa, T., Carninci, P.,  
Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F.,  
Imoto, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y.,  
Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M.,  
Nomura, K., Numazaki, R., Ohno, M., Oheato, N., Saito, R., Sakazume, N.,  
Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M.,  
Takeda, Y., Waki, K., Watanuki, A., Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Computational Analysis of Full-Length Mouse cDNAs Compared with  
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
10 (11), 1757-1771 (2000)  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics  
Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome  
Trust/MRC building Addenbrookes Hospital Cambridge) whose  
assistance we gratefully acknowledge.  
Please visit our web site (http://genome.gsc.riken.go.jp) for  
further details.

**FEATURES**  
source  
Location/Qualifiers  
1. .699  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="NOD"  
/db\_xref="taxon:10090"  
/clone="F630207H20"  
/cell\_type="NOD-derived CD11c +ve dendritic cells"  
/clone\_lib="RIKEN full-length enriched, NOD-derived CD11c  
+ve dendritic cells"

**ORIGIN**  
Alignment Scores:  
Pred. No.: 739 Length: 699  
Score: 44.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 87.50% Mismatches: 0  
Query Match: 88.00% Indels: 0  
DB: 6 Gaps: 0

US-10-014-101B-40 (1-9) x BY748808 (1-699)

**QY** 2 LeuGlyGlyLeuGlyGlnPheCys 9  
|||||:::|||||

**Db** 526 CTGGAGGGATGGGGCAGTTTGT 549

**RESULT 12**  
B1916220 846 bp mRNA linear EST 17-OCT-2001  
**LOCUS** B1916220 846 bp mRNA linear EST 17-OCT-2001  
**DEFINITION** 603178115F1 NIH\_MGC\_121 Homo sapiens cDNA clone IMAGE:5242539 5',

**ACCESSION**  
B1916220  
**VERSION**  
B1916220.1 GI:16199430  
**KEYWORDS**  
EST.  
**SOURCE**  
Homo sapiens (human)  
**ORGANISM**  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
**REFERENCE**  
1 (bases 1 to 846)  
**AUTHORS**  
NIH-MGC http://mgc.nci.nih.gov/  
**TITLE**  
National Institutes of Health, Mammalian Gene Collection (MGC)  
**JOURNAL**  
Unpublished (1999)  
**COMMENT**  
Contact: Robert Strausberg, Ph.D.  
Email: cgaps-femail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM11611 row: m column: 04  
High quality sequence stop: 239.

**FEATURES**  
source  
Location/Qualifiers  
1. .846  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5242539"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_121"  
/notes="Organ: brain; Vector: pCMV-SPORT6; Site: 1: NOL1;  
Site 2: EcoRV (destroyed); RNA source anonymous pool of 3  
fetal brains, female age 20 weeks, female age 24 weeks,  
and male age 26 weeks. Library is oligo-dT primed and  
directionally cloned (EcoRV site is destroyed upon  
cloning). Average insert size 1.7 kb, insert size range  
0.7-3.5 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 017. Note:  
this is a NIH\_MGC Library."

**ORIGIN**  
Alignment Scores:  
Pred. No.: 906 Length: 846  
Score: 44.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 87.50% Mismatches: 0  
Query Match: 88.00% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-40 (1-9) x B1916220 (1-846)

**QY** 2 LeuGlyGlyLeuGlyGlnPheCys 9  
|||||:::|||||

**Db** 542 CTAGGGGGTATAGGCCAATTCGTC 565

**RESULT 13**  
BQ439107/c 1007 bp mRNA linear EST 24-MAY-2002  
**LOCUS** BQ439107 1007 bp mRNA linear EST 24-MAY-2002  
**DEFINITION** AGENCOURT\_7786109 NIH\_MGC\_68 Homo sapiens cDNA clone IMAGE:6024189  
5', mRNA sequence.

**ACCESSION**  
BQ439107  
**VERSION**  
BQ439107.1 GI:21178183  
**KEYWORDS**  
EST.  
**SOURCE**  
Homo sapiens (human)  
**ORGANISM**  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
**REFERENCE**  
1 (bases 1 to 1007)  
**AUTHORS**  
NIH-MGC http://mgc.nci.nih.gov/  
**TITLE**  
National Institutes of Health, Mammalian Gene Collection (MGC)  
**JOURNAL**  
Unpublished (1999)

## COMMENT

Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-r@mail.nih.gov  
 Tissue Procurement: DCTD/DP/Gazdar  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLAM13234 row: e column: 22  
 High quality sequence stop: 484.  
 Location/Qualifiers  
 1. .1007  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6024189"  
 /tissue\_type="large cell carcinoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 68"  
 /notes="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 1.8 kb. Library constructed by Life  
 Technologies."

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.09e+03 Length: 1007  
 Score: 44.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 2  
 Best Local Similarity: 77.78% Mismatches: 0  
 Query Match: 88.00% Indels: 0  
 DB: 5 Gaps: 0

US-10-014-101B-40 (1-9) x BQ439107 (1-1007)

Qy 1 ValLeuGlyLeuGlyGlnPheCys 9

Db 711 GTTATTGGGGGCTCTGGCGCTTCTGCG 685

## RESULT 14

BB608644  
 LOCUS BB608644 RIKEN full-length enriched, 2 days pregnant adult female  
 DEFINITION ovary Mus musculus cDNA clone E330007A21 5', mRNA sequence.

ACCESSION BB608644.1 GI:11563821

VERSION EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 276)

## AUTHORS

Aizawa, K., Akahira, S., Akimura, T., Arai, A., Arakawa, T.,  
 Carninci, P., Hanagaki, T., Hayatsu, N., Hirose, T., Hirozane, T.,  
 Hodojima, Y., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kawai, J.,  
 Kojima, Y., Konno, H., Kusakabe, M., Matsuyama, T., Miyazaki, A.,  
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Okazaki, Y.,  
 Okido, T., Owa, C., Sakai, C., Sakai, K., Sasaki, D., Sato, K.,  
 Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,  
 Suzuki, H., Tagawa, A., Takahashi, F., Tanaka, T., Toya, T.,  
 Watahiki, A., Yamamura, T., Yaeunishi, A., Yoshida, K., Yoshihiki, A.,  
 Muramatsu, M. and Hayashizaki, Y.

## TITLE

RIKEN Mouse ESTs (Aizawa, K. et al. 2000)

## JOURNAL

Unpublished (2000)  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216

Email: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>

## FEATURES

## source

1. .276  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="E330007A21"  
 /sex="female"  
 /tissue\_type="ovary"  
 /dev\_stage="2 days pregnant adult"  
 /lab\_host="DH10B"  
 /clone\_lib="RIKEN full-length enriched, 2 days pregnant  
 adult female ovary"  
 /note="Site 1: SalI; Site 2: BamHI; cDNA library was  
 prepared and sequenced in Mouse Genome Encyclopedia  
 Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in  
 RIKEN. Division of Experimental Animal Research in Riken  
 contributed to prepare mouse tissues. 1st strand cDNA was  
 primed with a primer [5'  
 GAGAGAGAGCGCGCACTCGAGTTTTTTTTTTTNN 3'], cDNA was  
 prepared by using trehalose thermo-activated reverse  
 transcriptase and subsequently enriched for full-length by  
 cap-trapper. Second strand cDNA was prepared with the  
 primer adapter of sequence [5'  
 GAGAGAGAGATTCTCGAGTTAATTAATTCCTCCCTCCCCCCC 3']. cDNA  
 was cleaved with BamHI and XhoI. Vector: a modified  
 pBluescript KS(+) after bulk excision from Lambda FIC I."

## ORIGIN

Alignment Scores:  
 Pred. No.: 419 Length: 276  
 Score: 43.00 Matches: 7  
 Percent Similarity: 88.89% Conservative: 1  
 Best Local Similarity: 77.78% Mismatches: 1  
 Query Match: 86.00% Indels: 0  
 DB: 2 Gaps: 0

US-10-014-101B-40 (1-9) x BB608644 (1-276)

Qy 1 ValLeuGlyLeuGlyGlnPheCys 9

Db 225 GTCTTTGGGGGCTAGTGCAATTCGT 251

## RESULT 15

BB942814/c

LOCUS BB942814 370 bp mRNA linear EST 15-JAN-2003  
 DEFINITION ax30a01.x1 Hembase; Erythroid Progenitor Cells (LCB:ax library)  
 Homo sapiens cDNA clone ax30a01 random, mRNA sequence.

ACCESSION BB942814.1 GI:14342186

VERSION EST.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 370)

## AUTHORS

Gubin, A.N., Njoroge, J.M., Bouffard, G.G. and Miller, J.L.

## TITLE

Gene expression in proliferating human erythroid cells

JOURNAL  
MEDLINE  
PUBMED  
COMMENT

Genomics 59 (2), 168-177 (1999)  
99339981  
10409428  
Contact: Jeffery L. Miller  
Laboratory of Chemical Biology  
National Institute of Diabetes  
Building 10, Room 9B17, National Institutes of Health, Bethesda, MD  
20892, USA  
Tel: 301 402 2373  
Fax: 301 435 5148  
Email: jlm7f@nih.gov  
DNA Sequencing and analyses by National Institutes of Health  
Intramural Sequencing Center (NISC).  
Plate: 30 row: a column: 01  
Seq primer: -21M13 forward primer (ABI).  
Location/Qualifiers  
1. .370  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="ax30a01"  
/sex="unknown"  
/tissue\_type="blood"  
/cell\_type="Erythroid Cells"  
/cell\_line="Primary Culture of Peripheral Blood  
Mononuclear Cells"  
/dev\_stage="Progenitor; EPO responsive CD71+""  
/lab\_host="SOLR"  
/clone\_lib="Hembase; Erythroid Progenitor Cells (LCB:ax  
library)"  
/notes="Organ: blood; Vector: Lambda ZAP II; Site 1: EcoRI;  
Site 2: EcoRI; 65,000 proliferating erythroid cells from  
the buffy coat of a blood donation were obtained by flow  
cytometric separation after a 5-day culture period in the  
presence of erythropoietin. Total RNA was purified from  
the sorted cell population using TRIzol reagent. RNA (0.3  
ug) was converted into double stranded cDNA using  
Clontech's CapFinder cDNA Library Construction Kit  
(Clontech) according to the manufacturer's protocol and  
cloned into EcoRI digested Lambda Zap II vector  
(Stratagene). The phage library was amplified once prior  
to in vivo excision in SOLR cells. Individual colonies  
were grown, and the cDNA inserts were sequenced in high  
throughput (NIH intramural sequencing center  
<http://www.nisc.nih.gov/>)."

ORIGIN

Alignment Scores:  
Pred. No.: 573 Length: 370  
Score: 43.00 Matches: 8  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 1  
Query Match: 86.00% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-40 (1-9) x BG942814 (1-370)

QY 1 ValLeuGlyLeuGlyGlnPheCys 9  
|||||  
DB 75 GTTCTGGGTGGCTGGGGCTGTTCTGC 49  
|||||

RESULT 16  
AI552042  
LOCUS  
DEFINITION  
mm73g06.v1 Stratagene mouse macrophage (#937306) Mus musculus cDNA  
clone IMAGE:534106 5', mRNA sequence.  
ACCESSION  
AI552042  
VERSION  
AI552042.1 GI:4484405  
KEYWORDS  
EST.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
1 (bases 1 to 425)  
Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,  
Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y.,  
Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R.,  
Ritter, B., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,  
Waterston, R. and Wilson, R.  
The WashU-NCI Mouse EST Project 1999  
Unpublished (1999)  
Contact: Marra M/WashU-NCI Mouse EST Project 1999  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LML; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
This read is a RESEQUENCE of a previously sequenced mouse clone  
This read has been verified (found to hit its original self in the  
correct orientation)  
MGI:321042  
Seq primer: -40RP from Gibco  
High quality sequence stop: 395  
POLYA=No.

FEATURES  
source  
1. .425 Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:534106"  
/tissue\_type="macrophage"  
/lab\_host="SOLR (kanamycin resistant)"  
/clone\_lib="Stratagene mouse macrophage (#937306)"  
/note="Organ: blood; Vector: pBluescript SK-; Site 1:  
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:  
Oligo dt. WEHI-3 cell line. Average insert size: 1.5 kb;  
Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGCAG  
3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'."

ORIGIN

Alignment Scores:  
Pred. No.: 665 Length: 425  
Score: 43.00 Matches: 8  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 1  
Query Match: 86.00% Indels: 0  
DB: 1 Gaps: 0

US-10-014-101B-40 (1-9) x AI552042 (1-425)

QY 1 ValLeuGlyLeuGlyGlnPheCys 9  
|||||  
DB 280 GTACTGGGAGGCTGGGTTTGTTCG 306  
|||||

RESULT 17  
AA072623  
LOCUS  
DEFINITION  
mm73g06.r1 Stratagene mouse macrophage (#937306) Mus musculus cDNA  
clone IMAGE:534106 5', mRNA sequence.  
ACCESSION  
AA072623  
VERSION  
AA072623.1 GI:1594354  
KEYWORDS  
EST.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1 (bases 1 to 433)  
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,  
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,  
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and  
Waterston, R.  
The WashU-HHMI Mouse EST Project

JOURNAL  
COMMENT

Unpublished (1996)  
Contact: Marra M/Mouse EST Project  
WashU-HMI Mouse EST Project  
Washington University School  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:321042

Seq primer: -28ml3 rev1 ET from Amersham  
High quality sequence stop: 371.

## FEATURES

## source

1. .433  
Location/Qualifiers

/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:534106"  
/tissue\_type="macrophage"  
/dev\_stage="WEHI-3 Cell line"  
/lab\_host="SOLR (kanamycin resistant)"  
/clone\_lib="Stratagene mouse macrophage (#937306)"  
/note="Organ: blood; Vector: pBluescript SK-; Site 1:  
EcotRI; Site 2: XhoI; Cloned unidirectionally. Primer:  
Oligo dT. WEHI-3 cell line. Average insert size: 1.5 kb;  
Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGG  
3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3' "

## ORIGIN

## Alignment Scores:

Pred. No.: 679 Length: 433  
Score: 43.00 Matches: 8  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 1  
Query Match: 86.00% Indels: 0  
DB: 1 Gaps: 0

US-10-014-101B-40 (1-9) x AA072623 (1-433)

QY 1 ValLeuGlyLeuGlyGlnPheCys 9

Db 275 GTACTGGAGGCTTGGTTGTTTTC 301

## RESULT 18

CE667699

LOCUS

DEFINITION tigr-gss-dog-17000329256102 Dog Library Canis familiaris genomic,  
genomic survey sequence.

ACCESSION CE667699

VERSION CE667699.1 GI:36986567

KEYWORDS GSS.

SOURCE Canis familiaris (dog)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

1 (bases 1 to 443)

Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,

Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and

Venter,J.C.

The dog genome: survey sequencing and comparative analysis

Science 301 (5641), 1898-1903 (2003).

22875432

PUBMED 14512627

CONTACT: Kirkness EF

The Institute for Genomic Research

Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,

Rockville, MD 20850, USA

Tel: 301-838-0200

Fax: 301-838-0208

Email: ekirknes@tigr.org

Class: shotgun

Location/Qualifiers

## FEATURES

## source

1. .443  
/organism="Canis familiaris"  
/mol\_type="genomic DNA"  
/strain="Standard Poodle"  
/db\_xref="taxon:9615"  
/clone\_lib="Dog Library"  
/note="Site 1: BstXI; Libraries were prepared from  
peripheral blood"

## ORIGIN

## Alignment Scores:

Pred. No.: 696 Length: 443  
Score: 43.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 87.50% Mismatches: 0  
Query Match: 86.00% Indels: 0  
DB: 9 Gaps: 0

US-10-014-101B-40 (1-9) x CE667699 (1-443)

QY 2 LeuGlyGlyLeuGlyGlnPheCys 9

Db 358 GTGGGGGGTTCGCCAATTCGTC 381

## RESULT 19

AW180619/C

LOCUS

DEFINITION MGA0762f MGA Library Mycosphaerella graminicola cDNA clone MGA0762  
5' similar to PROBABLE LACTOYLGLUTATHIONE LYASE, mRNA sequence.

ACCESSION AW180619

VERSION AW180619.1 GI:6447814

KEYWORDS EST.

SOURCE Mycosphaerella graminicola

ORGANISM Mycosphaerella graminicola

Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes et

Chaetothyrriomycetes incertae sedis; Mycosphaerellaceae;

Mycosphaerella.

REFERENCE 1 (bases 1 to 498)

AUTHORS Keon,J.P.R., Bailey,A.M. and Hargreaves,J.A.

TITLE A group of expressed cDNA sequences from the wheat fungal leaf

blotch pathogen, Mycosphaerella graminicola (Septoria tritici)

JOURNAL Fungal Genet. Biol. 29 (2), 118-133 (2000)

MEDLINE 20374020

PUBMED 10919380

## COMMENT

Contact: Hargreaves JA

Cell Biology Department

IACR-Long Ashton Research Station

Long Ashton, Bristol, BS41 9AF, UK

Tel: +44(0)1275 392181

Fax: +44(0)1275 394281

Email: john.hargreaves@bbsrc.ac.uk

Insert Length: 1100 Std Error: 0.00

Seq primer: M13 reverse

Location/Qualifiers

## FEATURES

## source

1. .498  
/organism="Mycosphaerella graminicola"  
/mol\_type="mRNA"  
/strain="Strit"  
/db\_xref="taxon:54734"  
/clone="MGA0762"  
/clone\_lib="Mga Library"  
/note="Vector: pSPORT1; Library constructed from cultures  
utilizing ammonium ions as a source of nitrogen"

## ORIGIN

## Alignment Scores:

Pred. No.: 788 Length: 498  
Score: 43.00 Matches: 7  
Percent Similarity: 88.89% Conservative: 1  
Best Local Similarity: 77.78% Mismatches: 1  
Query Match: 86.00% Indels: 0  
DB: 2 Gaps: 0

US-10-014-101B-40 (1-9) x AW180619 (1-498)

Qy 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
|||||  
Db 108 GTACTCGAGGATCGGTCTATTTTGC 82

RESULT 20  
CD217252 516 bp mRNA linear EST 20-MAY-2003  
DEFINITION TGESTzy122b05.y1 TgMAS Tachyzoite cDNA Library Toxoplasma gondii  
cDNA clone TGESTzy122b05.y1 5' similar to SW:RL39\_HUMAN P02404 60S  
RIBOSOMAL PROTEIN L39. ;, mRNA sequence.

ACCESSION CD217252  
VERSION CD217252.1 GI:30957225  
KEYWORDS EST.

SOURCE Toxoplasma gondii

ORGANISM Toxoplasma gondii

REFERENCE 1 (bases 1 to 516)

AUTHORS Tang,K., Cole,R., Fogarty,S., Sibley,L.D., Ajioke,J.A., White,M.,

Clifton,S., Pape,D., Martin,J., Wylie,T., Dante,M., Marra,M.,

Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons,M.,

Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R., Ronko,I.,

Kennedy,S., Maguire,L., Waterston,R. and Wilson,R.

Toxoplasma EST Project

Unpublished (2001)

CONTACT: Clifton, S.

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: toxo@watson.wustl.edu

Contact David Sibley (toxest@borcim.wustl.edu) for further

information relating to organism, libraries, or clone availability.

Possible reversed clone: similarity on wrong strand

Seq primer: -40RP from Gibco.

Location/Qualifiers

1. 516

/organism="Toxoplasma gondii"

/mol\_type="mRNA"

/strain="Tachyzoite"

/db\_xref="taxon:5811"

/clone="TGESTzy122b05.y1"

/dev\_stage="Tachyzoite"

/lab\_host="Electroten Blue cells (Stratagene)"

/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:

XhoI; The cDNA library was constructed by Keliang Tang,

and Robert Cole at Washington University. cDNA was

synthesized from Poly(A)+ mRNA using an oligo-d(T) primer

containing a XhoI site. Following second strand synthesis,

EcoRI adapters were ligated to the cDNA, and products were

size-selected on sephacryl S500. The cDNA were

directionally cloned into the EcoRI/XhoI prepared

pBluescript II SK+ vector, and electroporated into

Electroten Blue cells (Stratagene). The library may

contain a small percentage of host or bacterial

contaminants."

# ORIGIN

Alignment Scores:  
Pred. No.: 819 Length: 516  
Score: 43.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 3  
Best Local Similarity: 66.67% Mismatches: 0  
Query Match: 86.00% Indels: 0  
DB: 6 Gaps: 0

US-10-014-101B-40 (1-9) x CD217252 (1-516)

Qy 1 ValLeuGlyGlyLeuGlyGlnPheCys 9

Db 431 ATCCTCGGGGATCGGCGGTCTGC 457

RESULT 21

LOCUS BU699322

DEFINITION

BU699322 517 bp mRNA linear EST 09-OCT-2002

musculus cDNA 5', mRNA sequence.

ACCESSION BU699322

VERSION BU699322.1 GI:23615910

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 517)

AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Phillips,R.L., Ernst,R.E., Brunk,B.P., Ivanova,N., Mahan,M.A.,

Deanehan,J.K., Moore,K.A., Overton,G.C. and Lemischka,I.R.

The genetic program of hematopoietic stem cells

Science 288 (5471), 1635-1640 (2000)

10834841

20295303

COMMENT

Contact: Lemischka, Ihor R.

Department of Molecular Biology

Princeton University

Lewis Thomas Laboratory, Washington Road, Princeton, NJ 08544, USA

Tel: 609 258 2838

Fax: 609 258 2759

Email: ilemischka@molbio.princeton.edu

These ESTs are derived from a subtracted cDNA library enriched for

stem products expressed in day 14-14.5 fetal liver hematopoietic

stem cells defined as Lineageneg/lo, AA4.lpos, ckitpos, Ly6A/E

(Sca-1)pos

Seq primer: M13Reverse or T7.

Location/Qualifiers

1. 517

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/tissue type="Hematopoietic"

/cell\_type="Stem Cells; Lineageneg/lo, AA4.lpos, ckitpos,

Ly6A/E (Sca-1)pos"

/dev\_stage="Embryonic day 14-14.5"

/lab\_host="DH10B"

/clone lib="Hematopoietic Stem Cell Subtracted Library"

/notes="Organ: Fetal Liver; Vector: Sport 1; Site 1: Sal I;

Site 2: Not 1; Two directionally cloned cDNA libraries

were made from fetal liver hematopoietic stem cells

enriched to be Lineageneg/lo, AA4.lpos, ckitpos, Ly6A/E

(Sca-1)pos called Scapos and from AA4.lneg fetal liver

cells. Subtractive hybridization was performed by

hybridization of the target, Scapos, single stranded cDNA

library in pSport1 to biotinylated RNA transcribed from

the driver, AA4.lneg cDNA library in pSport2 with inserts

cloned in the complementary orientation. For detailed

protocols and additional information please see our

website at http://stemcell.princeton.edu."

# ORIGIN

Alignment Scores:  
Pred. No.: 821 Length: 517  
Score: 43.00 Matches: 8  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 1  
Query Match: 86.00% Indels: 0  
DB: 5 Gaps: 0

US-10-014-101B-40 (1-9) x BU699322 (1-517)

Qy 1 ValLeuGlyGlyLeuGlyGlnPheCys 9

Db 111 GTACTCGAGGCTTGGTTGTTTGC 137

```

RESULT 22
BI848950/c
LOCUS      530 bp      mRNA      linear      EST 04-OCT-2001
DEFINITION 471549 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION  BI848950
VERSION     BI848950.1  GI:15961469
KEYWORDS    EST.
SOURCE      Bos taurus (cow)
ORGANISM    Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
            Bovinae; Bos.
REFERENCE   1 (bases 1 to 530)
AUTHORS    Smith,T.P.L., Grose,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
            Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C.,
            Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A.,
            Chitko-McKown,C.G., Perte,G., Holt,I., Karamycheva,S., Liang,F.,
            Quackenbush,J. and Keefe,J.W.
TITLE      Sequence evaluation of four pooled-tissue normalized bovine cDNA
            libraries and construction of a gene index for cattle
JOURNAL    Genome Res. 11 (4), 626-630 (2001)
MEDLINE    21180013
PUBMED     11282978
COMMENT    Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smitht@mail.marc.usda.gov
            Single pass sequencing. Bases called and alt trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTCCAGTCACGACG
Plate: 99 row: A column: 8
Seq primer: ATTTAGGTGACACTATAG.
FEATURES   source
            1..530
            /organism="Bos taurus"
            /mol_type="mRNA"
            /db_xref="taxon:9913"
            /tissue_type="pooled"
            /lab_host="DH10B"
            /clone_lib="MARC 2BOV"
            /note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
            Library made from pooled tissue from testis, thymus,
            semitendinosus muscle, longissimus muscle, pancreas,
            adrenal, and endometrium."
ORIGIN
Alignment Scores:
Pred. No.:      843      Length:      530
Score:          43.00    Matches:      7
Percent Similarity: 88.89%      Conservative: 1
Best Local Similarity: 77.78%      Mismatches: 1
Query Match:      86.00%      Indels:      0
DB:              4          Gaps:      0
US-10-014-101b-40 (1-9) x BI848950 (1-530)
Qy      1 ValLeuGlyGlyLeuGlyGlnPheCys 9
Db      335 ATCTTGGAGCCCTGGCTGCTGCTGCT 309
RESULT 23
BX643028/c
LOCUS      576 bp      mRNA      linear      EST 04-SEP-2003
DEFINITION DKFZp781G0822 r1 781 (synonym: hlcc4) Homo sapiens cDNA clone
ACCESSION  BX643028
VERSION     DKFZp781G0822.5', mRNA sequence.
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 576)
AUTHORS    Koehler,K., Beyer,A., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
            Fobo,G., Han,M. and Wiemann,S.
TITLE      EST (Koehler,K., Beyer,A., Mewes,H.W., Weil,B., Amid,C., et al.)
JOURNAL    Unpublished (2003)
COMMENT    Contact: MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by BMFZ (Biomedical Research Center at the Heinrich-
            Heine-University, Duesseldorf/Germany) within the cDNA sequencing
            consortium of the German Genome Project. No sl sequence available.
            This clone (DKFZp781G0822) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES   source
            1..576
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="DKFZp781G0822"
            /dev_stage="adult"
            /lab_host="DH10B"
            /clone_lib="781 (synonym: hlcc4)"
            /note="Vector: pSPORT1_Sfi; Site_1: SfiIA; Site_2: SfiIB;
            cDNA-collection"
ORIGIN
Alignment Scores:
Pred. No.:      921      Length:      576
Score:          43.00    Matches:      8
Percent Similarity: 88.89%      Conservative: 0
Best Local Similarity: 88.89%      Mismatches: 1
Query Match:      86.00%      Indels:      0
DB:              5          Gaps:      0
US-10-014-101b-40 (1-9) x BX643028 (1-576)
Qy      1 ValLeuGlyGlyLeuGlyGlnPheCys 9
Db      103 GTTCTGGGTGGCTGGGCTGTTCTGC 77
RESULT 24
BP768145/c
LOCUS      594 bp      mRNA      linear      EST 10-JUL-2004
DEFINITION BP768145 mouse (C57BL/6) pancreatic islet library with
            recombination-based method Mus musculus cDNA clone mid30071 3',
            mRNA sequence.
ACCESSION  BP768145
VERSION     BP768145.1  GI:50226843
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 594)
            Nishimura,M., Yokoi,N., Miki,T., Horikawa,Y., Yoshioka,H.,
            Takeda,J., Ohara,O. and Seino,S.
            Construction of a multi-functional cDNA library specific for mouse
            pancreatic islets and its application to microarray
            Unpublished (2004)
            Contact: Susumu Seino
            Division of Cellular and Molecular Medicine
            Kobe University Graduate School of Medicine
            7-5-1 Kusunoki-cho, Chuo-ku, Kobe, Hyogo 650-0017, Japan
            Tel: 81-78-382-5360

```





Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2=RC3-HT0585-010  
 400-023-h11et3=2000-04-01&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 45  
 High quality sequence stop: 685.  
 Location/Qualifiers

## FEATURES

source

1. 714  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="HT0585"  
 /note="Organ: head neck; Vector: puc18; Site 1: SmaI;  
 Site 2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the pUC 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.16e+03 Length: 714  
 Score: 43.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 87.50% Mismatches: 0  
 Query Match: 86.00% Indels: 0  
 DB: 2 Gaps: 0

US-10-014-101B-40 (1-9) x BE176609 (1-714)

QY 2 LeuGlyGlyLeuGlyGlnPheCys 9

Db 57 CTAGGAGGAGTAGGCCAGTCTTC 34

## RESULT 29

LOCUS CL714569 827 bp DNA linear GSS 26-JUL-2004  
 DEFINITION OR\_BBa0039F11.r OR\_BBa Oryza rufipogon genomic clone OR\_BBa0039F11  
 3', genomic survey sequence.

ACCESSION CL714569

VERSION CL714569.1 GI:50601607

KEYWORDS GSS.

SOURCE Oryza rufipogon

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE 1 (bases 1 to 827)

AUTHORS Kim,H., Yu,X., Stum,D., Yost,D., Rao,K., Luo,M., Jetty,R.,  
 Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.

TITLE ONAP Project

JOURNAL Unpublished (2004)

COMMENT Contact: Rod A. Wing

Arizona Genomics Institute

University of Arizona

Forbes Building Room 303, Tucson, AZ 85721-0036, USA

Tel: 520 626 9595

Fax: 520 621 1259

Email: http://genome.arizona.edu

PCR Primers

FORWARD: TAA TAC GAC TCA CTA TAG GG

BACKWARD: CAC TCA TTA GGC ACC CCA

Insert Length: 161 Std Error: 0.00

Plate: 0039 row: F column: 11

Seq primer: CAC TCA TTA GGC ACC CCA

Class: BAC ends.

Location/Qualifiers

1. 827

## FEATURES

source

/organism="Oryza rufipogon"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:4529"  
 /clone="OR\_BBa0039F11"  
 /tissue\_type="young leaves"  
 /lab\_host="DH10B-T1 phage resistant"  
 /clone\_lib="OR\_BBa"  
 /note="Vector: pAGIBAC1; Site\_1: HindIII; Site\_2: HindIII"

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.36e+03 Length: 827  
 Score: 43.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 2  
 Best Local Similarity: 77.78% Mismatches: 0  
 Query Match: 86.00% Indels: 0  
 DB: 9 Gaps: 0

US-10-014-101B-40 (1-9) x CL714569 (1-827)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9

Db 459 GTCTAGGAGGTCGGAAGGTTTCT 485

## RESULT 30

LOCUS BQ050317

DEFINITION BQ050317 1143 bp mRNA linear EST 29-MAR-2002  
 AGENCOURT 7050735 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:5784242  
 5', mRNA sequence.

ACCESSION BQ050317

VERSION BQ050317.1 GI:19809657

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1143)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

CNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM12869 row: h column: 03

High quality sequence stop: 525.

## FEATURES

source

1. 1143  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5784242"  
 /tissue\_type="leiomyosarcoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 71"  
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: Sali; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 2.1 kb."

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.92e+03 Length: 1143  
 Score: 43.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 87.50% Mismatches: 0  
 Query Match: 86.00% Indels: 0  
 DB: 5 Gaps: 0



19-44 (1999)  
Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

**FEATURES**  
source  
1. .253  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="D830007B12"  
/tissue\_type="heart"  
/dev\_stage="16 days neonate"  
/lab\_host="DH10B"  
/clone\_lib="RIKEN full-length enriched, 16 days neonate heart"  
/notes="Site 1: Sall; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'-GAGAGAGAGCGCGCACTCGAGTTTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGAGAGATTCGAGTTAATAATTAATCCCGCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

**ORIGIN**  
Alignment Scores:  
Pred. No.: 586 Length: 253  
Score: 42.00 Matches: 6  
Percent Similarity: 88.89% Conservatives: 2  
Best Local Similarity: 66.67% Mismatches: 1  
Query Match: 84.00% Indels: 0  
DB: 2 Gaps: 0

US-10-014-101B-40 (1-9) x BB603301 (1-253)  
QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
:::|||||  
Db 191 ATCTTGGGGGTTAGTGAAATCTGT 217

**RESULT 33**  
AL361639  
LOCUS  
DEFINITION AL361639 ICRFP 522 and 523 Mus musculus cDNA clone M9129A60 5', mRNA sequence.  
EST.  
ACCESSION AL361639  
VERSION AL361639.1 GI:9692561  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE Eickhoff, H., Schuchhardt, J., Ivanov, I., Meier-Ewert, S., O'Brien, J., Malik, A., Tandon, N., Woleki, E., Rohlf, E., Nyarsik, L., Reinhardt, R., Niefeld, W. and Lehrach, H.  
Tissue gene expression analysis using arrayed normalized cDNA libraries  
Genome Res. (2000) In press  
Contact: MPIMG  
Abt. Lehrach  
Max Planck Institut Fuer Molekulare Genetik  
Innestraße 73, Berlin, 14195 Germany  
The cDNA libraries ICRFP 522 and 523 were normalized with oligonucleotide fingerprinting, resulting in a unique subset of 5376 cDNA clones.

**FEATURES**  
source  
1. .400  
Location/Qualifiers

/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="M9129A60"  
/tissue\_type="embryo"  
/dev\_stage="9 and 12 pc embryo"  
/clone\_lib="ICRFP 522 and 523"

**ORIGIN**  
Alignment Scores:  
Pred. No.: 957 Length: 400  
Score: 42.00 Matches: 6  
Percent Similarity: 88.89% Conservatives: 2  
Best Local Similarity: 66.67% Mismatches: 1  
Query Match: 84.00% Indels: 0  
DB: 1 Gaps: 0

US-10-014-101B-40 (1-9) x AL361639 (1-400)  
QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9  
:::|||||  
Db 310 ATAATCGAGGATTGGCAGTTTCTGT 336

**RESULT 34**  
AL361639  
LOCUS  
DEFINITION AL361639 SP 0042 Al C10 SP6E Strongylocentrotus purpuratus, purple sea urchin, sperm genomic BAC library Strongylocentrotus purpuratus genomic clone Plate=42 Col=19 Row=E, genomic survey sequence.  
ACCESSION AL361639  
VERSION AL361639.1 GI:8293272  
KEYWORDS  
SOURCE Strongylocentrotus purpuratus  
ORGANISM  
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa; Echinoida; Euechinoida; Echinacea; Echinoida; Strongylocentrotidae; Strongylocentrotus.  
REFERENCE 1 (bases 1 to 413)  
AUTHORS Cameron, R.A., Mahairas, G., Rast, J.P., Martinez, P., Biondi, T.R., Swartzell, S., Wallace, J.C., Poustka, A.J., Livingston, B.T., Wray, G.A., Ettensohn, C.A., Lehrach, H., Britten, R.J., Davidson, E.H. and Hood, L.  
A sea urchin genome project: Sequence scan, virtual map, and additional resources  
Proc. Natl. Acad. Sci. U.S.A. 97 (17), 9514-9518 (2000)  
PUBMED 10920195  
COMMENT Contact: Cameron, RA, Davidson, EH, Hood, L  
Division of Biology 156-29  
California Institute of Technology  
Pasadena California 91125, USA  
Tel: (626) 395-8421  
Fax: (626) 793-3047  
Email: acameron@caltech.edu  
Plate: 42 row: E column: 19  
Seq primer: SP6  
Class: BAC ends  
High quality sequence stop: 413.  
Location/Qualifiers  
1. .413  
/organism="Strongylocentrotus purpuratus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7668"  
/clone="plate=42 Col=19 Row=E"  
/clone\_lib="Strongylocentrotus purpuratus, purple sea urchin, sperm genomic BAC library"  
/note="Organ: sperm; Vector: BAC3.6; BAC Clones in E-Coli DH10B"

**ORIGIN**  
Alignment Scores:  
Pred. No.: 990 Length: 413





Fax: 81-75-705-1113  
Email: yutaka@ascidian.zool.kyoto-u.ac.jp.

## FEATURES

## source

1. .446  
/organism="Ciona savignyi"  
/mol\_type="mRNA"  
/db\_xref="taxon:51511"  
/clone="csga085g17"  
/dev\_stage="gastrula stage"  
/clone\_lib="Yutaka Satou unpublished cDNA library (csga)"

## ORIGIN

Alignment Scores:  
Pred. No.: 1.08e+03 Length: 446  
Score: 42.00 Matches: 7  
Percent Similarity: 88.89% Conservative: 1  
Best Local Similarity: 77.78% Mismatches: 1  
Query Match: 84.00% Indels: 0  
DB: 5 Gaps: 0

US-10-014-101B-40 (1-9) x BW523847 (1-446)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9

Db 69 GTCTTGGAGGATTGGAGGTTTGT 43

## RESULT 39

## CN605793

LOCUS  
DEFINITION CN605793 452 bp mRNA linear EST 30-SEP-2004  
PEL0290 Primordial expressed cDNA library lambda ZAP Lentinula edodes cDNA clone PEL0290 3' similar to 60S ribosomal protein L30 (S. pombe), mRNA sequence.

ACCESSION CN605793

VERSION CN605793.1 GI:52852731

KEYWORDS EST.

SOURCE Lentinula edodes (shiitake mushroom)

## ORGANISM

Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes; Agaricales; Tricholomataceae; Lentinula.

REFERENCE 1 (bases 1 to 452)

AUTHORS Ng, T. P., Bian, X. L., Ng, W. L. and Kwan, H. S.

TITLE Expressed Sequence Tags from Lentinula edodes

JOURNAL Unpublished (2004)

## COMMENT

Contact: Kwan, H. S.

Food and Nutritional Sciences Programme, Molecular Biotechnology

Programme, Department of Biology

The Chinese University of Hong Kong

Science Centre, Shatin, N.T., Hong Kong SAR

Tel: 852-26096285

Fax: 852-26035745

Email: hoishankwan@cuhk.edu.hk

PCR Primers

FORWARD: T3

BACKWARD: T7

Insert length: 452 Std Error: 0.00

Seq primer: Oligo dt25N

POLYA-No.

## FEATURES

## source

1. .452  
/organism="Lentinula edodes"  
/mol\_type="mRNA"  
/strain="L54"  
/db\_xref="taxon:5353"  
/clone="PEL0290"  
/sex="Dikaryon"  
/dev\_stage="Primordium"  
/clone\_lib="Primordial expressed cDNA library lambda ZAP"

/notes="Vector: lambda ZAP Express(TM) vector (Stratagene); Site 1: EcoRI; Site 2: EcoRI; Vegetative mycelium strain L54 was cultivated in sawdust medium. Primordium was stimulated by induction of cold shock. mRNA from primordium was extracted. The subsequently synthesized cDNA library was cloned into ZAP Express(TM)

## ORIGIN

Alignment Scores:  
Pred. No.: 1.09e+03 Length: 452  
Score: 42.00 Matches: 7  
Percent Similarity: 88.89% Conservative: 1  
Best Local Similarity: 77.78% Mismatches: 1  
Query Match: 84.00% Indels: 0  
DB: 7 Gaps: 0

US-10-014-101B-40 (1-9) x CN605793 (1-452)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9

Db 385 ATCTTGGGGCGCTTGGACTTTTGT 411

## RESULT 40

## CN6072048/c

LOCUS  
DEFINITION CN6072048 476 bp mRNA linear EST 15-JUN-2004  
GR\_Ea31A03.f GR\_Ea Gossypium raimondii cDNA clone GR\_Ea31A03 5', mRNA sequence.

ACCESSION CN6072048

VERSION CN6072048.1 GI:48741529

KEYWORDS EST.

SOURCE Gossypium raimondii

## ORGANISM

Gossypium raimondii  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.

REFERENCE 1 (bases 1 to 476)

## AUTHORS

Kim, H., Yu, Y., Rudrna, D., Hatfield, J., Stum, D., Mueller, C.,

Udall, J. A., Rapp, R. A., Wendel, J. F., Rao, K., Soderlund, C. and

Wing, R. A.

Global assembly of Cotton ESTs

## TITLE

## JOURNAL

## COMMENT

Contact: Rod A. Wing

Arizona Genomics Institute

The University of Arizona

Forbes Building Room 303, Tucson, AZ, 85721-0036, USA

Tel: 520 626 9595

Fax: 520 621 1259

Email: http://genome.arizona.edu

Plate: 31 row: A column: 03.

## FEATURES

## source

1. .476  
/organism="Gossypium raimondii"  
/mol\_type="mRNA"  
/db\_xref="taxon:29730"  
/clone="GR\_Ea31A03"  
/tissue\_type="whole seedlings"  
/dev\_stage="first true leaves"  
/lab\_host="DH108"  
/clone\_lib="GR\_Ea"  
/note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2: EcoRV; Library made by Invitrogen with RNA supplied by Wendle lab. Directional cloned into NotI-EV. Colonies plated/picked by AGI. More glycerol clones held in -80."

## ORIGIN

## Alignment Scores:

Pred. No.: 1.15e+03 Length: 476  
Score: 42.00 Matches: 6  
Percent Similarity: 88.89% Conservative: 2  
Best Local Similarity: 66.67% Mismatches: 1  
Query Match: 84.00% Indels: 0  
DB: 7 Gaps: 0

US-10-014-101B-40 (1-9) x CN6072048 (1-476)

QY 1 ValLeuGlyGlyLeuGlyGlnPheCys 9

Db 387 ATCATGGGGCGCTTGGGCAATTTTTC 361

Search completed: February 18, 2005, 07:00:58  
Job time : 1069.28 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 17, 2005, 21:59:48 ; Search time 822.72 Seconds  
(without alignments)  
353.378 Million cell updates/sec

Title: US-10-014-101B-41  
Perfect score: 27  
Sequence: 1 ITRARI 6

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues  
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DRV=xlh  
-O=/cgn2 1/USPTO.spool/US10014101/runat.16022005.075845.7936/app.query.fasta\_1.796  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=40 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10014101 @CGN 1.1 2886 @runat.16022005.075845.7936 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length DB ID	Description
C 1	27	100.0	111 12 SYNMRNA03	M22020 Synthetic 0
2	27	100.0	130 11 BV079618	BV079618 E2M2-130
3	27	100.0	151 8 HVU234763	AJ234763 Hordeum v
4	27	100.0	159 6 AX935409	AX935409 Sequence

C 5	27	100.0	192 6 CQ648448	CQ648448 Sequence
C 6	27	100.0	195 6 CQ651934	CQ651934 Sequence
C 7	27	100.0	195 6 AX607583	AX607583 Sequence
C 8	27	100.0	372 8 AF490285	AF490285 Phytolphth
C 9	27	100.0	372 8 AF490286	AF490286 Phytolphth
C 10	27	100.0	393 3 AF408671	AF408671 Trichogra
C 11	27	100.0	395 3 AF408673	AF408673 Trichogra
C 12	27	100.0	397 3 AF408672	AF408672 Trichogra
C 13	27	100.0	488 11 GS8417	GS8417 SHGC-104567
C 14	27	100.0	493 6 CQ476090	CQ476090 Sequence
C 15	27	100.0	540 9 HSA332323	HSA332323 Homo sapi
C 16	27	100.0	546 8 AY191142	AY191142 Oryza sat
C 17	27	100.0	547 8 AY191137	AY191137 Pyricularia
C 18	27	100.0	555 11 AF096559	AF096559 Rattus no
C 19	27	100.0	601 11 BV192133	BV192133 sqmml7346
C 20	27	100.0	648 4 AF232017	AF232017 Tamandua
C 21	27	100.0	700 3 AY330210	AY330210 Toxoplasma
C 22	27	100.0	724 9 HSA327616	HSA327616 Homo sapi
C 23	27	100.0	730 3 AY330204	AY330204 Neospora
C 24	27	100.0	800 4 AJ606422	AJ606422 Marmosops
C 25	27	100.0	805 8 AF534679	AF534679 Pyricularia
C 26	27	100.0	816 8 ZMU82201	ZMU82201 Zea mays pa
C 27	27	100.0	860 11 BV065128	BV065128 S212P6051
C 28	27	100.0	872 8 AK103395	AK103395 Oryza sat
C 29	27	100.0	874 6 BD275463	BD275463 47 Human
C 30	27	100.0	887 11 CNS06ELJ	CNS06ELJ T7 end of
C 31	27	100.0	900 6 E16279	E16279 gDNA encodi
C 32	27	100.0	984 4 AF434177	AF434177 Thylanya
C 33	27	100.0	984 6 AX935448	AX935448 Sequence
C 34	27	100.0	985 4 AF434176	AF434176 Thylanya
C 35	27	100.0	987 6 AX935453	AX935453 Sequence
C 36	27	100.0	987 6 AX935455	AX935455 Sequence
C 37	27	100.0	988 4 AF431925	AF431925 Thylanya
C 38	27	100.0	990 4 AF431923	AF431923 Thylanya
C 39	27	100.0	1039 6 CQ497388	CQ497388 Sequence
C 40	27	100.0	1045 3 AY333784	AY333784 Toxoplasma
C 41	27	100.0	1059 13 AY487468	AY487468 Unculture
C 42	27	100.0	1089 3 AF459563	AF459563 Drosophila
C 43	27	100.0	1098 3 AF459564	AF459564 Drosophila
C 44	27	100.0	1099 3 AF459545	AF459545 Drosophila
C 45	27	100.0	1099 3 AF459546	AF459546 Drosophila

ALIGNMENTS

RESULT 1	SYNMRNA03/c	SYNMRNA03	111 bp	mrna	linear	SYN 27-APR-1993
LOCUS	SYNMRNA03	Synthetic 003 mrna.				
DEFINITION	M22020					
ACCESSION	M22020.1	GI:208873				
VERSION		artificial gene.				
KEYWORDS		synthetic construct				
SOURCE		synthetic construct				
ORGANISM		other sequences; artificial sequences.				
REFERENCE	1 (bases 1 to 111)					
AUTHORS	Calogero,R.A., Pon,C.L., Canonaco,M.A. and Gualerzi,C.O.					
TITLE	Selection of the mRNA translation initiation region by Escherichia coli ribosomes					
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 85 (17), 6427-6431 (1988)					
MEDLINE	88320453					
PUBMED	3045916					
COMMENT	Original					
FEATURES	Location/Qualifiers					
source	1..111					
	/organism="synthetic construct"					
	/mol_type="mrna"					
	/db_xref="taxon:32630"					
mrna	1..111					
	/product="003 mrna"					
CDS	25..57					
	/note="003 mrna peptide"					
	/codon_start=1					

/transl\_table=11  
 /protein\_id="AAA72693.1"  
 /db\_xref="GI:208874"  
 /translation="MFTITITPFT"

## ORIGIN

Alignment Scores: Length: 111  
 Pred. No.: 253 Matches: 6  
 Score: 27.00  
 Percent Similarity: 100.00%  
 Best Local Similarity: 100.00%  
 Query Match: 100.00%  
 DB: 12 Gaps: 0

US-10-014-101B-41 (1-6) x SYNMRNA03 (1-111)

QY 1 IleThrArgAlaArgile 6

Db 21 ATACCAAGGCGCGAATT 4

## RESULT 2

BV079618 130 bp DNA linear STS 20-JUL-2004  
 LOCUS E2M2-130 Hessian fly genomic DNA Mayetiola destructor STS genomic,  
 DEFINITION sequence tagged site.  
 ACCESSION BV079618  
 VERSION BV079618.1 GI:34787367  
 KEYWORDS STS.  
 SOURCE Mayetiola destructor (Hessian fly)  
 ORGANISM Mayetiola destructor

REFERENCE 1 (bases 1 to 130)  
 AUTHORS Behura,S.K., Valiente,F.H., Rider,S.D. Jr., Shun-Chen,M.,  
 Jackson,S. and Stuart,J.J.  
 TITLE A physically anchored genetic map and linkage to avirulence reveals  
 recombination suppression over the proximal region of hessian fly  
 chromosome A2  
 JOURNAL Genetics 167 (1), 343-355 (2004)  
 PUBMED 15166159

COMMENT Contact: Jeff Stuart  
 Department of Entomology  
 Purdue University  
 901 W State St., West Lafayette, IN 47907, USA  
 Primer A: None provided  
 Primer B: None provided  
 Protocol:

Template: 20-30ng  
 Primer: each 20 pmoles  
 dNTPs: each 200um  
 Taq Pol: 0.1 units/ul  
 Total Vol: 25 ul

## Buffer:

MgCl2: 2.5mM  
 KCl: 50 mM  
 Tris-Cl: 10mM  
 pH: 8.3  
 Location/Qualifiers

## FEATURES

source  
 1..130  
 /organism="Mayetiola destructor"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:39758"  
 /clone\_lib="Hessian fly genomic DNA"  
 <1..>130

## STS

## ORIGIN

Alignment Scores: Length: 130  
 Pred. No.: 294 Matches: 6  
 Score: 27.00  
 Percent Similarity: 100.00%  
 Query Match: 100.00%  
 DB: 6 Gaps: 0

Best Local Similarity: 100.00%  
 Query Match: 100.00%  
 DB: 11 Gaps: 0

US-10-014-101B-41 (1-6) x BV079618 (1-130)

QY 1 IleThrArgAlaArgile 6

Db 36 ATACCAAGGCGCTAGAATA 53

## RESULT 3

HVU234763 151 bp DNA linear PLN 07-JAN-2000  
 LOCUS Hordeum vulgare genomic DNA fragment; clone MMG2081.rev.  
 DEFINITION  
 ACCESSION AJ234763  
 VERSION AJ234763.1 GI:3819546  
 KEYWORDS RFLP marker.  
 SOURCE Hordeum vulgare subsp. vulgare  
 ORGANISM Hordeum vulgare subsp. vulgare

REFERENCE 1 (bases 1 to 151)  
 AUTHORS Michalek,W., Kunzel,G. and Graner,A.  
 TITLE Sequence analysis and gene identification in a set of mapped RFLP  
 markers in barley (Hordeum vulgare)  
 JOURNAL Genome 42 (5), 849-853 (1999)  
 MEDLINE 20051696  
 PUBMED 10584307

REFERENCE 2 (bases 1 to 151)

AUTHORS Michalek,W.  
 TITLE Direct Submission  
 JOURNAL Submitted (29-OCT-1998) Michalek W., Institute of Plant Genetics  
 and Crop Plant Research (IPK), Corrensstr. 3, 06466 Gatersleben,  
 Germany  
 FEATURES Location/Qualifiers  
 source  
 1..151  
 /organism="Hordeum vulgare subsp. vulgare"  
 /mol\_type="genomic DNA"  
 /cultivar="Igri"  
 /sub\_species="vulgare"  
 /db\_xref="taxon:112509"  
 /clone="MMG2081.rev"  
 /tissue\_type="leaf"  
 misc\_feature  
 1..151  
 /note="vector: pBluescript; Primer: M13rev"

## ORIGIN

Alignment Scores: Length: 151  
 Pred. No.: 340 Matches: 6  
 Score: 27.00  
 Percent Similarity: 100.00%  
 Best Local Similarity: 100.00%  
 Query Match: 100.00%  
 DB: 8 Gaps: 0

US-10-014-101B-41 (1-6) x HVU234763 (1-151)

QY 1 IleThrArgAlaArgile 6

Db 87 ATACCAAGGCGCAAGATT 104

## RESULT 4

AX935409 159 bp DNA linear PAT 05-JAN-2004  
 LOCUS Sequence 13 from Patent WO03089475.  
 DEFINITION  
 ACCESSION AX935409  
 VERSION AX935409.1 GI:40642182  
 KEYWORDS Raphanus sativus (radish)  
 SOURCE Raphanus sativus  
 ORGANISM Raphanus sativus

REFERENCE 1 (bases 1 to 151)  
 AUTHORS Michalek,W., Kunzel,G. and Graner,A.  
 TITLE Sequence analysis and gene identification in a set of mapped RFLP  
 markers in barley (Hordeum vulgare)  
 JOURNAL Genome 42 (5), 849-853 (1999)  
 MEDLINE 20051696  
 PUBMED 10584307

REFERENCE 2 (bases 1 to 151)  
 AUTHORS Michalek,W.  
 TITLE Direct Submission  
 JOURNAL Submitted (29-OCT-1998) Michalek W., Institute of Plant Genetics  
 and Crop Plant Research (IPK), Corrensstr. 3, 06466 Gatersleben,  
 Germany  
 FEATURES Location/Qualifiers  
 source  
 1..151  
 /organism="Hordeum vulgare subsp. vulgare"  
 /mol\_type="genomic DNA"  
 /cultivar="Igri"  
 /sub\_species="vulgare"  
 /db\_xref="taxon:112509"  
 /clone="MMG2081.rev"  
 /tissue\_type="leaf"  
 misc\_feature  
 1..151  
 /note="vector: pBluescript; Primer: M13rev"

REFERENCE 1 (bases 1 to 151)  
 AUTHORS Michalek,W., Kunzel,G. and Graner,A.  
 TITLE Sequence analysis and gene identification in a set of mapped RFLP  
 markers in barley (Hordeum vulgare)  
 JOURNAL Genome 42 (5), 849-853 (1999)  
 MEDLINE 20051696  
 PUBMED 10584307

REFERENCE 2 (bases 1 to 151)  
 AUTHORS Michalek,W.  
 TITLE Direct Submission  
 JOURNAL Submitted (29-OCT-1998) Michalek W., Institute of Plant Genetics  
 and Crop Plant Research (IPK), Corrensstr. 3, 06466 Gatersleben,  
 Germany  
 FEATURES Location/Qualifiers  
 source  
 1..151  
 /organism="Hordeum vulgare subsp. vulgare"  
 /mol\_type="genomic DNA"  
 /cultivar="Igri"  
 /sub\_species="vulgare"  
 /db\_xref="taxon:112509"  
 /clone="MMG2081.rev"  
 /tissue\_type="leaf"  
 misc\_feature  
 1..151  
 /note="vector: pBluescript; Primer: M13rev"

rosids; eurosids II; Brassicales; Brassicaceae; Raphanus.

REFERENCE 1  
AUTHORS Peschen,D., Fischer,R., Schillberg,S., Liao,Y.C. and Dorfmueller,S.  
TITLE Antibodies, recombinant antibodies, recombinant antibody fragments  
and fusions mediated plant disease resistance against fungi  
JOURNAL Patent: WO 03089475-A 13 30-OCT-2003;  
Fraunhofer-Gesellschaft zur Foerderung der angewandten Forschung  
e.V. (DE)

FEATURES  
source Location/Qualifiers  
1. .159  
/organism="Raphanus sativus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3726"

ORIGIN  
Alignment Scores: 357 Length: 159  
Pred. No.: 27.00 Matches: 6  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 100.00% Gaps: 0  
DB: 6

US-10-014-101B-41 (1-6) x AX935409 (1-159)

QY 1 IleThrArgAlaArgIle 6  
|||||  
DB 59 ATAACGCGTCAAGAAATC 76

RESULT 5  
LOCUS CQ648448/c 192 bp DNA linear PAT 02-FEB-2004  
DEFINITION Sequence 5405 from Patent W00234771.  
ACCESSION CQ648448  
VERSION CQ648448.1 GI:41683950  
KEYWORDS Streptococcus agalactiae  
SOURCE Streptococcus agalactiae  
ORGANISM Streptococcus agalactiae  
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

REFERENCE 1  
AUTHORS Telford,J., Massignani,V., Margarit y Ros,I., Grandi,G., Fraser,C.  
and Tettelin,H.  
TITLE Nucleic acids and proteins from streptococcus groups a & b  
JOURNAL Patent: WO 0234771-A 5405 02-MAY-2002;  
Chiron S.p.A. (IT) ; THE INSTITUTE FOR GENOMIC RESEARCH (US)

FEATURES  
source Location/Qualifiers  
1. .192  
/organism="Streptococcus agalactiae"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:1311"

ORIGIN  
Alignment Scores: 429 Length: 192  
Pred. No.: 27.00 Matches: 6  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 100.00% Gaps: 0  
DB: 6

US-10-014-101B-41 (1-6) x CQ648448 (1-192)

QY 1 IleThrArgAlaArgIle 6  
|||||  
DB 116 ATAACAAGAGCAGAAATT 99

RESULT 6  
LOCUS CQ651934/c 195 bp DNA linear PAT 02-FEB-2004  
DEFINITION Sequence 8891 from Patent W00234771.  
ACCESSION CQ651934  
VERSION CQ651934.1 GI:41685756

KEYWORDS Streptococcus agalactiae  
SOURCE Streptococcus agalactiae  
ORGANISM Streptococcus agalactiae  
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

REFERENCE 1  
AUTHORS Telford,J., Massignani,V., Margarit y Ros,I., Grandi,G., Fraser,C.  
and Tettelin,H.  
TITLE Nucleic acids and proteins from streptococcus groups a & b  
JOURNAL Patent: WO 0234771-A 8891 02-MAY-2002;  
Chiron S.p.A. (IT) ; THE INSTITUTE FOR GENOMIC RESEARCH (US)

FEATURES  
source Location/Qualifiers  
1. .195  
/organism="Streptococcus agalactiae"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:1311"

ORIGIN  
Alignment Scores: 435 Length: 195  
Pred. No.: 27.00 Matches: 6  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 100.00% Gaps: 0  
DB: 6

US-10-014-101B-41 (1-6) x CQ651934 (1-195)

QY 1 IleThrArgAlaArgIle 6  
|||||  
DB 116 ATAACAAGAGCAGAAATT 99

RESULT 7  
LOCUS AX607583/c 195 bp DNA linear PAT 17-FEB-2003  
DEFINITION Sequence 5512 from Patent W002092818.  
ACCESSION AX607583  
VERSION AX607583.1 GI:28403115  
KEYWORDS Streptococcus agalactiae  
SOURCE Streptococcus agalactiae  
ORGANISM Streptococcus agalactiae  
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

REFERENCE 1  
AUTHORS Glaeser,P., Rusniok,C., Chevalier,F., Frangeul,L., Lalioui,L.,  
Zouine,M., Couve,E., Buchrieser,C., Poyart,C., Trieu-Cuot,P. and  
Kunst,F.  
TITLE Streptococcus agalactiae genome sequence, use for developing  
vaccines, diagnostic tools, and for identifying therapeutic targets  
JOURNAL Patent: WO 02092818-A 5512 21-NOV-2002;  
INSTITUT PASTEUR (FR) ; CENTRE NATIONAL DE LA RECHERCHE  
SCIENTIFIQUE (CNRS) (FR)

FEATURES  
source Location/Qualifiers  
1. .195  
/organism="Streptococcus agalactiae"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:1311"

ORIGIN  
Alignment Scores: 435 Length: 195  
Pred. No.: 27.00 Matches: 6  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 100.00% Gaps: 0  
DB: 6

US-10-014-101B-41 (1-6) x AX607583 (1-195)

QY 1 IleThrArgAlaArgIle 6  
|||||  
DB 116 ATAACAAGAGCAGAAATT 99

```

RESULT 8
AF490285/c
LOCUS      372 bp      DNA      linear      PLN 26-JUL-2002
DEFINITION Phytophthora mirabilis strain 3010 clone 1 gypsy-like
            retrotransposon, partial sequence.
ACCESSION  AF490285
VERSION     AF490285.1 GI:20386310
KEYWORDS   Phytophthora mirabilis
SOURCE     Phytophthora mirabilis
ORGANISM   Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
            Phytophthora.
REFERENCE  1 (bases 1 to 372)
            Judelson,H.S.
            Sequence Variation and Genomic Amplification of a Family of
            Gypsy-like Elements in the Oomycete Genus Phytophthora
            Mol. Biol. Evol. 19 (8), 1313-1322 (2002)
JOURNAL    22135996
MEDLINE    12140243
PUBMED     12140243
REFERENCE  2 (bases 1 to 372)
            Judelson,H.S.
            Direct Submission
            Submitted (01-MAR-2002) Dept. of Plant Pathology, University of
            California, Webber Hall, Riverside, CA 92521, USA
COMMENT    Contains probable non-functional element.
FEATURES   Location/Qualifiers
            1..372
                /organism="Phytophthora mirabilis"
                /mol_type="genomic DNA"
                /strain="3010"
                /db_xref="taxon:129356"
                /clone="1"
            repeat_region <1..>372
                /transposon="gypsy-like retrotransposon"

ORIGIN
Alignment Scores:
Pred. No.:      812      Length:      372
Score:          27.00    Matches:      6
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:     100.00%  Indels:      0
DB:              8      Gaps:          0

US-10-014-101B-41 (1-6) x AF490285 (1-372)
QY      1 lleThrArgAlaArgIle 6
DB      206 ATTACGCGGCAAGAATC 189

RESULT 9
AF490286/c
LOCUS      372 bp      DNA      linear      PLN 26-JUL-2002
DEFINITION Phytophthora mirabilis strain 3010 clone 2 gypsy-like
            retrotransposon, partial sequence.
ACCESSION  AF490286
VERSION     AF490286.1 GI:20386311
KEYWORDS   Phytophthora mirabilis
SOURCE     Phytophthora mirabilis
ORGANISM   Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
            Phytophthora.
REFERENCE  1 (bases 1 to 372)
            Judelson,H.S.
            Sequence Variation and Genomic Amplification of a Family of
            Gypsy-like Elements in the Oomycete Genus Phytophthora
            Mol. Biol. Evol. 19 (8), 1313-1322 (2002)
JOURNAL    22135996
MEDLINE    12140243
PUBMED     12140243
REFERENCE  2 (bases 1 to 372)
            Judelson,H.S.
            Direct Submission
            Submitted (01-MAR-2002) Dept. of Plant Pathology, University of
            California, Webber Hall, Riverside, CA 92521, USA
COMMENT    Contains probable non-functional element.
FEATURES   Location/Qualifiers
            1..372
                /organism="Phytophthora mirabilis"
                /mol_type="genomic DNA"
                /strain="3010"
                /db_xref="taxon:129356"
                /clone="1"
            repeat_region <1..>372
                /transposon="gypsy-like retrotransposon"

ORIGIN
Alignment Scores:
Pred. No.:      812      Length:      372
Score:          27.00    Matches:      6
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:     100.00%  Indels:      0
DB:              8      Gaps:          0

US-10-014-101B-41 (1-6) x AF490285 (1-372)
QY      1 lleThrArgAlaArgIle 6
DB      206 ATTACGCGGCAAGAATC 189

RESULT 9
AF490286/c
LOCUS      372 bp      DNA      linear      PLN 26-JUL-2002
DEFINITION Phytophthora mirabilis strain 3010 clone 2 gypsy-like
            retrotransposon, partial sequence.
ACCESSION  AF490286
VERSION     AF490286.1 GI:20386311
KEYWORDS   Phytophthora mirabilis
SOURCE     Phytophthora mirabilis
ORGANISM   Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
            Phytophthora.
REFERENCE  1 (bases 1 to 372)
            Judelson,H.S.
            Sequence Variation and Genomic Amplification of a Family of
            Gypsy-like Elements in the Oomycete Genus Phytophthora
            Mol. Biol. Evol. 19 (8), 1313-1322 (2002)
JOURNAL    22135996
MEDLINE    12140243
PUBMED     12140243
REFERENCE  2 (bases 1 to 372)
            Judelson,H.S.
            Direct Submission
            Submitted (01-MAR-2002) Dept. of Plant Pathology, University of
            California, Webber Hall, Riverside, CA 92521, USA
COMMENT    Contains probable non-functional element.
FEATURES   Location/Qualifiers
            1..372
                /organism="Phytophthora mirabilis"
                /mol_type="genomic DNA"
                /strain="3010"
                /db_xref="taxon:129356"
                /clone="1"
            repeat_region <1..>372
                /transposon="gypsy-like retrotransposon"

ORIGIN
Alignment Scores:
Pred. No.:      812      Length:      372
Score:          27.00    Matches:      6
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:     100.00%  Indels:      0
DB:              8      Gaps:          0

US-10-014-101B-41 (1-6) x AF490286 (1-372)
QY      1 lleThrArgAlaArgIle 6
DB      206 ATTACGCGGCAAGAATC 189

RESULT 10
AF408671/c
LOCUS      393 bp      DNA      linear      INV 18-FEB-2002
DEFINITION Trichogramma alpha strain ACLE internal transcribed spacer 2,
            complete sequence.
ACCESSION  AF408671
VERSION     AF408671.1 GI:18699085
KEYWORDS   Trichogramma alpha
SOURCE     Trichogramma alpha
ORGANISM   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Hymenoptera; Apocrita; Chalcidoidea;
            Trichogrammatidae; Trichogramma.
REFERENCE  1 (bases 1 to 393)
            Pinto,J.D., Koopmanschap,A.B., Platner,G.R. and Stouthamer,R.
            The North American Trichogramma (Hymenoptera: Trichogrammatidae)
            parasitizing certain Tortricidae (Lepidoptera) on apple and pear,
            with ITS2 DNA characterizations and description of a new species
            Biol. Control 23, 134-142 (2002)
JOURNAL    2 (bases 1 to 393)
            Pinto,J.D., Koopmanschap,A.B., Platner,G.R. and Stouthamer,R.
            Direct Submission
            Submitted (10-AUG-2001) Entomology, University of California,
            Riverside, CA 92521, USA
COMMENT    Contains probable non-functional element.
FEATURES   Location/Qualifiers
            1..393
                /organism="Trichogramma alpha"
                /mol_type="genomic DNA"
                /strain="ACLE"
                /db_xref="taxon:185869"
            misc_RNA
                1..393
                /product="internal transcribed spacer 2"

ORIGIN
Alignment Scores:
Pred. No.:      857      Length:      393
Score:          27.00    Matches:      6
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:     100.00%  Indels:      0
DB:              3      Gaps:          0

US-10-014-101B-41 (1-6) x AF408671 (1-393)
QY      1 lleThrArgAlaArgIle 6

```

```

Db      334 ATACGGCGCGGCATT 317

RESULT 11
AF408673/c
LOCUS   395 bp      DNA      linear      INV 18-FEB-2002
DEFINITION Trichogramma alpha strain ADRK internal transcribed spacer 2,
complete sequence.
ACCESSION AF408673
VERSION   AF408673.1 GI:18699087
KEYWORDS
SOURCE
ORGANISM Trichogramma alpha
Trichogramma alpha
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Hymenoptera; Apocrita; Chalcidoidea;
Trichogrammatidae; Trichogramma.
REFERENCE 1 (bases 1 to 395)
AUTHORS Pinto, J.D., Koopmanschap, A.B., Platner, G.R. and Stouthamer, R.
TITLE The North American Trichogramma (Hymenoptera: Trichogrammatidae)
parasitizing certain Tortricidae (Lepidoptera) on apple and pear,
with ITS2 DNA characterizations and description of a new species
JOURNAL Biol. Control 23, 134-142 (2002)
REFERENCE 2 (bases 1 to 395)
AUTHORS Pinto, J.D., Koopmanschap, A.B., Platner, G.R. and Stouthamer, R.
TITLE Direct Submission
JOURNAL Submitted (10-AUG-2001) Entomology, University of California,
Riverside, CA 92521, USA

FEATURES
source
Location/Qualifiers
1..395
/organism="Trichogramma alpha"
/mol_type="genomic DNA"
/strain="ADRK"
/db_xref="taxon:185869"
misc_RNA
1..395
/product="internal transcribed spacer 2"

ORIGIN
Alignment Scores:
Pred. No.: 861 Length: 395
Score: 27.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-014-101B-41 (1-6) x AF408673 (1-395)

QY 1 IleThrArgAlaArgile 6
|||||
Db 336 ATACGGCGCGGCATT 319

RESULT 12
G58417/c
LOCUS   397 bp      DNA      linear      INV 18-FEB-2002
DEFINITION Trichogramma alpha strain ACOL internal transcribed spacer 2,
complete sequence.
ACCESSION AF408672
VERSION   AF408672.1 GI:18699086
KEYWORDS
SOURCE
ORGANISM Trichogramma alpha
Trichogramma alpha
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Hymenoptera; Apocrita; Chalcidoidea;
Trichogrammatidae; Trichogramma.
REFERENCE 1 (bases 1 to 397)
AUTHORS Pinto, J.D., Koopmanschap, A.B., Platner, G.R. and Stouthamer, R.
TITLE The North American Trichogramma (Hymenoptera: Trichogrammatidae)
parasitizing certain Tortricidae (Lepidoptera) on apple and pear,
with ITS2 DNA characterizations and description of a new species
JOURNAL Biol. Control 23, 134-142 (2002)
REFERENCE 2 (bases 1 to 397)
AUTHORS Pinto, J.D., Koopmanschap, A.B., Platner, G.R. and Stouthamer, R.
TITLE Direct Submission
JOURNAL Submitted (10-AUG-2001) Entomology, University of California,
Riverside, CA 92521, USA

FEATURES
source
Location/Qualifiers
1..397
/organism="Trichogramma alpha"
/mol_type="genomic DNA"
/strain="ACOL"
/db_xref="taxon:185869"
misc_RNA
1..397
/product="internal transcribed spacer 2"

ORIGIN
Alignment Scores:
Pred. No.: 865 Length: 397
Score: 27.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-014-101B-41 (1-6) x AF408672 (1-397)

QY 1 IleThrArgAlaArgile 6
|||||
Db 334 ATACGGCGCGGCATT 317

RESULT 13
G58417/c
LOCUS   488 bp      DNA      linear      STS 30-MAR-2000
DEFINITION SHGC-104567 Human Homo sapiens STS genomic, sequence tagged site.
ACCESSION G58417
VERSION   G58417.1 GI:6123736
KEYWORDS STS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Homo sapiens
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 488)
AUTHORS Olivier, M. and Cox, D.R.
TITLE Unpublished, Olivier, M., Cox, D.R. (2000)
JOURNAL Unpublished (2000)
COMMENT
Contact: Michael Olivier, David R. Cox
Stanford Human Genome Center
Stanford University School of Medicine
4005 Miranda Ave. 2nd Fl., Palo Alto, CA 94025, USA
Tel: (650) 320-5800
Fax: (650) 320-5801
Email: olivier@hgc.stanford.edu
Primer A: TCGAATCTGATGGCTCAATAAA
Primer B: CCTGGTTTGGATGGAACCTATTT
STS size: 322
PCR Profile:
Initial incubation: 95 degrees C for 10 minutes
Denaturation: 94 degrees C for 30 seconds
Annealing: 60 degrees C for 30 seconds
Polymerization: 72 degrees C for 23 seconds
PCR Cycles: 30
Thermal Cycler: Perkin Elmer 9700

Protocol:
Template: 25 ng
Primer: each 1 uM
dNTPs: each 200 uM
AmpliTaq Gold Polymerase: 0.07 units/ul
Total Vol: 5 ul

Buffer: MgCl2: 2.5 mM
KCl: 50 mM
Tris-HCl: 10 mM
pH: 8.3

BAC ends sequenced at TIGR from the RPC111 BAC library. Designed
and developed at the Stanford Human Genome Center.

```

```

FEATURES
  source
    Location/Qualifiers
      1..488
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /map="2"
        /clone_lib="Human"
      18..339
        primer_bind
      18..41
        primer_bind
      complement(317..339)
ORIGIN
  Alignment Scores:
    Pred. No.: 1..06e+03 Length: 488
    Score: 27.00 Matches: 6
    Percent Similarity: 100.00% Conservative: 0
    Best Local Similarity: 100.00% Mismatches: 0
    Query Match: 100.00% Indels: 0
    DB: 11 Gaps: 0
  US-10-014-101B-41 (1-6) x G58417 (1-488)
  QY 1 IleThrArgAlaArgile 6
  Db 315 ATTACTAGACTAGGATT 298
  RESULT 14
  LOCUS CO476090 493 bp DNA linear PAT 30-JAN-2004
  DEFINITION Sequence 7957 from Patent WO0160860.
  ACCESSION CO476090
  VERSION CO476090.1 GI:41441709
  KEYWORDS
  SOURCE Homo sapiens (human)
  ORGANISM
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE 1
  AUTHORS Schlegel,R., Endege,W.O. and Monahan,J.E.
  TITLE Genes differentially expressed in human prostate cancer and their
  JOURNAL use
  PATENT Patent: WO 0160860-A 7957 23-AUG-2001;
  Millennium Predictive Medicine, Inc. (US)
  FEATURES
    source
      Location/Qualifiers
        1..493
          /organism="Homo sapiens"
          /mol_type="unassigned DNA"
          /db_xref="taxon:9606"
ORIGIN
  Alignment Scores:
    Pred. No.: 1..07e+03 Length: 493
    Score: 27.00 Matches: 6
    Percent Similarity: 100.00% Conservative: 0
    Best Local Similarity: 100.00% Mismatches: 0
    Query Match: 100.00% Indels: 0
    DB: 6 Gaps: 0
  US-10-014-101B-41 (1-6) x CQ476090 (1-493)
  QY 1 IleThrArgAlaArgile 6
  Db 98 ATTACCGTCGCCGATA 115
  RESULT 15
  LOCUS HSA332323 540 bp DNA linear PRI 18-JUL-2002
  DEFINITION Homo sapiens genomic sequence surrounding NotI site, clone
  H5J-DO1RS.
  ACCESSION AJ332323
  VERSION AJ332323.1 GI:15876741
  KEYWORDS
  SOURCE Homo sapiens (human)
  ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE 1
  AUTHORS Kutsenko,A.S., Gizatullin,R.Z., Al-Amin,A.N., Wang,F., Kvasha,S.M.,
    Podowski,R.M., Matushkin,Y.G., Gyanchandani,A., Muravenko,O.V.,
    Levitsky,V.G., Kolchanov,N.A., Protopopov,A.I., Kashuba,V.I.,
    Kisselev,L.L., Wasserman,W., Wahlestedt,C. and Zabarovsky,E.R.
  TITLE NotI flanking sequences: a tool for gene discovery and verification
    of the human genome
  JOURNAL Nucleic Acids Res. 30 (14), 3163-3170 (2002)
  MEDLINE 22131767
  PUBMED 12136098
  REFERENCE 2 (bases 1 to 540)
  AUTHORS Zabarovsky,E.R.
  TITLE Direct Submission
  JOURNAL Submitted (16-MAY-2001) Microbiology and Tumorbiology Centre,
    Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77,
    Sweden
  FEATURES
    source
      Location/Qualifiers
        1..540
          /organism="Homo sapiens"
          /mol_type="genomic DNA"
          /db_xref="taxon:9606"
          /clone="HSJ-DO1RS"
ORIGIN
  Alignment Scores:
    Pred. No.: 1..16e+03 Length: 540
    Score: 27.00 Matches: 6
    Percent Similarity: 100.00% Conservative: 0
    Best Local Similarity: 100.00% Mismatches: 0
    Query Match: 100.00% Indels: 0
    DB: 9 Gaps: 0
  US-10-014-101B-41 (1-6) x HSA332323 (1-540)
  QY 1 IleThrArgAlaArgile 6
  Db 294 ATCACCGTCGCCGATC 311
  RESULT 16
  LOCUS AY191142 546 bp DNA linear PLN 01-JUN-2004
  DEFINITION Osyris wightiana trna-Leu (trnL) gene, partial sequence;
    chloroplast gene for chloroplast product.
  ACCESSION AY191142
  VERSION AY191142.1 GI:37779727
  KEYWORDS
  SOURCE chloroplast Osyris wightiana
  ORGANISM
    Osyris wightiana
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
    Santalales; Santalaceae; Osyris.
  REFERENCE 1 (bases 1 to 546)
  AUTHORS Han,R.L., Hao,G. and Zhang,D.X.
  TITLE Interfamilial relationships as revealed by chloroplast trnL intron
    sequences
  JOURNAL Unpublished
  REFERENCE 2 (bases 1 to 546)
  AUTHORS Han,R.L., Hao,G. and Zhang,D.X.
  TITLE Direct Submission
  JOURNAL Submitted (05-DEC-2002) Centre for Systematic and Evolutionary
    Botany, South China Institute of Botany, Leyiju, Guangzhou,
    Guangdong 510650, P.R. China
  FEATURES
    source
      Location/Qualifiers
        1..546
          /organism="Osyris wightiana"
          /organelle="plastid:chloroplast"
          /mol_type="genomic DNA"
          /db_xref="taxon:210363"
          <1..>546
          /gene="trnL"
    gene

```

```

intron      <1..514
            /gene="trnL"
tRNA        <515..>546
            /gene="trnL"
            /product="tRNA-Leu"

ORIGIN
Alignment Scores:
Pred. No.: 1.18e+03      Length: 546
Score: 27.00             Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 8                    Gaps: 0

US-10-014-101B-41 (1-6) x AY191142 (1-546)

QY 1 IleThrArgAlaArgile 6
Db 449 ATTACTCGGCGGAGAATA 466

RESULT 17
LOCUS AY191137
DEFINITION Pyralia sinensis tRNA-Leu (trnL) gene, partial sequence;
            chloroplast gene for chloroplast product.
ACCESSION AY191137
VERSION AY191137.1 GI:37779722
KEYWORDS chloroplast Pyralia sinensis
SOURCE Pyralia sinensis
ORGANISM Pyralia sinensis
REFERENCE 1 (bases 1 to 547)
AUTHORS Han,R.L., Hao,G. and Zhang,D.X.
TITLE Interfamilial relationships as revealed by chloroplast trnL intron
        sequences
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 547)
AUTHORS Han,R.L., Hao,G. and Zhang,D.X.
TITLE Direct Submission
JOURNAL Submitted (05-DEC-2002) Centre for Systematic and Evolutionary
        Botany, South China Institute of Botany, Leyifu, Guangzhou,
        Guangdong 510650, P.R. China
FEATURES
source 1..547
        /organism="Pyralia sinensis"
        /organelle="plastid:chloroplast"
        /mol_type="genomic DNA"
        /db_xref="taxon:227912"
gene <1..>547
        /gene="trnL"
intron <1..531
        /gene="trnL"
tRNA <532..>547
        /gene="trnL"
        /product="tRNA-Leu"

ORIGIN
Alignment Scores:
Pred. No.: 1.18e+03      Length: 547
Score: 27.00             Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 8                    Gaps: 0

US-10-014-101B-41 (1-6) x AY191137 (1-547)

QY 1 IleThrArgAlaArgile 6
Db 456 ATTACTCGGCGGAGAATA 473

```

---

```

RESULT 18
LOCUS AF096559
DEFINITION Rattus norvegicus clone D5Uwm49, sequence tagged site.
ACCESSION AF096559
VERSION AF096559.1 GI:5565747
KEYWORDS STS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
REFERENCE 1 (bases 1 to 555)
AUTHORS Lan,H., Shepel,L.A., Haag,J.D. and Gould,M.N.
TITLE Linkage mapping of rat chromosome 5 markers generated from
        chromosome-specific libraries
JOURNAL Mamm. Genome 10 (7), 687-691 (1999)
MEDLINE 99315326
PUBMED 10384040
REFERENCE 2 (bases 1 to 555)
AUTHORS Lan,H., Shepel,L.A., Haag,J.D. and Gould,M.N.
TITLE Direct Submission
JOURNAL Submitted (02-OCT-1998) McCardle Laboratory for Cancer Research,
        University of Wisconsin Medical School, 1400 University Avenue,
        Madison, WI 53706, USA
FEATURES
source 1..555
        /organism="Rattus norvegicus"
        /mol_type="genomic DNA"
        /strain="Copenhagen"
        /db_xref="taxon:10116"
        /chromosome="5"
        /clone="D5Uwm49"
        primer_bind 215..235
        primer_bind complement(446..463)
ORIGIN
Alignment Scores:
Pred. No.: 1.2e+03      Length: 555
Score: 27.00             Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 11                   Gaps: 0

US-10-014-101B-41 (1-6) x AF096559 (1-555)

QY 1 IleThrArgAlaArgile 6
Db 78 ATAACCTAGAGCTAGGATT 95

RESULT 19
LOCUS BV192133/c
DEFINITION sqm173469 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
        tagged site.
ACCESSION BV192133
VERSION BV192133.1 GI:48034424
KEYWORDS STS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 601)
AUTHORS Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
        Cantor,C.R. and Braun,A.
TITLE Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
        Regions
JOURNAL Genome Res. (2004) In press
COMMENT Contact: Andreas Braun
        Pharmaceuticals division

```

Sequenom, Inc.  
3595 John Hopkins Court, San Diego, CA 92121, USA  
Tel: 18582029018  
Fax: 18582029020  
Email: abra@sequenom.com  
Primer A: No primer sequence submitted  
Primer B: No primer sequence submitted  
STS size: 601.

## FEATURES

source  
1. .601  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone\_lib="Human DNA (Sequenom)"  
<1. .>601

## STS

ORIGIN

Alignment Scores:  
Pred. No.: 1.29e+03 Length: 601  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 11 Gaps: 0

US-10-014-101B-41 (1-6) x BV192133 (1-601)

QY 1 lleThrArgAlaArgIle 6  
|||||  
Db 482 ATAACTAGGCGCTAGGATT 465

## RESULT 20

AF232017/c

LOCUS AF232017 648 bp DNA linear MAM 27-FEB-2001  
DEFINITION Tamandua tetradactyla clone 1 cytochrome b gene, partial cds;  
mitochondrial gene for mitochondrial product.

ACCESSION AF232017

VERSION AF232017.1 GI:13128899

## KEYWORDS

SOURCE mitochondrion Tamandua tetradactyla (southern tamandua)

## ORGANISM

Tamandua tetradactyla  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Edentata; Myrmecophagidae; Tamandua.  
1 (bases 1 to 648)  
Greenwood, A.D., Castresana, J., Feldmaier-Fuchs, G. and Paabo, S.  
A molecular phylogeny of two extinct sloths  
Mol. Phylogenet. Evol. 18 (1), 94-103 (2001)  
21097350  
MEDLINE 11161746  
PUBMED

REFERENCE 2 (bases 1 to 648)

Greenwood, A.D., Castresana, J., Feldmaier-Fuchs, G. and Paabo, S.

Direct Submission

Submitted (08-FEB-2000) Vertebrate Zoology, American Museum of

Natural History, Central Park West at 79th Street, New York, NY

10024, USA

## FEATURES

source  
1. .648  
/organism="Tamandua tetradactyla"  
/organelle="mitochondrion"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:48850"  
/clone="1"  
1. .>648  
/codon\_start=1  
/transl\_table=2  
/product="cytochrome b"  
/protein\_id="AAK13065.1"  
/db\_xref="GI:13128890"

## CDS

translation="MTHIRKTHPLFKIINQSFIDLPAPNISAWNFGSLGLICLIHQ  
ILTGLFLAHYTSDDITAFSSVTHICRDVNYGWLIRYIHASGASMEFMCILYLVHVRGL  
YVGSGLYLETWNIGVILLATWATTFMGVYLPWGQMSFWGATVITNLLSAIPYIGTDL  
VENWINGFSVDKATLRFPPAFHPLFPILLALVMTLLFLHETGNSNPLGLSSAMD"

## ORIGIN

## Alignment Scores:

Pred. No.: 1.39e+03 Length: 648  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-41 (1-6) x AF232017 (1-648)

QY 1 lleThrArgAlaArgIle 6  
|||||  
Db 581 ATTACTAGGCGCTAGGATA 564

## RESULT 21

AY330210

LOCUS AY330210 700 bp DNA linear INV 07-JUN-2004  
DEFINITION Toxoplasma gondii clone Tgsat350-3 satellite DNA sequence.

ACCESSION AY330210

VERSION AY330210.1 GI:33243869

## KEYWORDS

SOURCE Toxoplasma gondii

ORGANISM Toxoplasma gondii

REFERENCE 1 (bases 1 to 700)

Sarcocystidae; Toxoplasma.

Clemente, M., De Miguel, N., Lia, V.V., Matrajt, M. and Angel, S.O.

Structure Analysis of Two Toxoplasma gondii and Neospora caninum

Satellite DNA Families and Evolution of Their Common Monomeric

Sequence

J. Mol. Evol. 58 (5), 557-567 (2004)

PUBMED 15170259

REFERENCE 2 (bases 1 to 700)

Clemente, M., de Miguel, N. and Angel, S.O.

Direct Submission

Submitted (24-JUN-2003) Parasitologia Molecular, IIB-INTSICH, Camino

de Circunvalacion Laguna Km 6, Chascomus, Buenos Aires B713011WA,

Argentina

FEATURES

Location/Qualifiers

source 1. .700

/organism="Toxoplasma gondii"

/mol\_type="genomic DNA"

/db\_xref="taxon:5811"

/clone="Tgsat350-3"

repeat\_region 1. .700

/note="satellite"

## ORIGIN

## Alignment Scores:

Pred. No.: 1.5e+03 Length: 700  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-10-014-101B-41 (1-6) x AY330210 (1-700)

QY 1 lleThrArgAlaArgIle 6  
|||||  
Db 428 ATCACTCGGCGCAGGATT 445

## RESULT 22

HS3327616

LOCUS HS3327616 724 bp DNA linear PRI 18-JUL-2002

DEFINITION Homo sapiens genomic sequence surrounding NotI site, clone

NR1-CF6R.

ACCESSION AJ327616

VERSION AJ327616.1 GI:15872034

## KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo..
1 (bases 1 to 724)
Podsenko,A.S., Gizatullin,R.Z., Al-Amin,A.N., Wang,F., Kvaasha,S.M.,
Kutsenko,R.S., Matushkin,Y.G., Gyanchandani,A., Muravenko,O.V.,
Levitsky,V.G., Kolchanov,N.A., Protodopov,A.I., Kashuba,V.I.,
Kiselev,I.L., Wasserman,W., Wahlstedt,C. and Zabarovsky,E.R.
Noti flanking sequences: a tool for gene discovery and verification
of the human genome
Nucleic Acids Res. 30 (14), 3163-3170 (2002)
2131767
PUBMED
12136098
2 (bases 1 to 724)
Zabarovsky,E.R.
Direct Submission
Submitted (16-MAY-2001) Microbiology and Tumorbiology Centre,
Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77,
Sweden
FEATURES             Location/Qualifiers
     source            1..724
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /clone="NR1-CF6R"
ORIGIN
Alignment Scores:
Pred. No.:           1.55e+03           Length:           724
Score:               27.00              Matches:           6
Percent Similarity: 100.00%             Conservative:      0
Best Local Similarity: 100.00%           Mismatches:       0
Query Match:         100.00%             Indels:           0
DB:                  9                  Gaps:             0

US-10-014-101B-41 (1-6) x HSA327616 (1-724)

Qy  1 IleThrArgAlaArglie 6
    |||||
Db  172 ATTACCGGGCAGGTATA 189

RESULT 23
AY330204
LOCUS      AY330204
DEFINITION Neospora caninum clone NcSat350-2 satellite DNA sequence.
ACCESSION AY330204
VERSION    AY330204.1 GI:33243692
KEYWORDS
SOURCE     Neospora caninum
ORGANISM   Neospora caninum
            Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
            Sarcocystidae; Neospora.
1 (bases 1 to 730)
Clemente,M., De Miguel,N., Lia,V.V., Matrajt,M. and Angel,S.O.
Structure Analysis of Two Topoplasma gondii and Neospora caninum
Satellite DNA Families and Evolution of Their Common Monomeric
Sequence
J Mol Evol. 58 (5), 557-567 (2004)
15170259
2 (bases 1 to 730)
Clemente,M., de Miguel,N. and Angel,S.O.
Direct Submission
Submitted (23-JUN-2003) Parasitologia Molecular, IIB-INTECH, Camino
de Circunvalacion Laguna Km 6, Chascomus, Buenos Aires B713011WA,
Argentina
FEATURES             Location/Qualifiers
     source            1..730
                        /organism="Neospora caninum"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:29176"
                        /clone="NcSat350-2"
                        /note="satellite"
repeat_region        1..730
ORIGIN

```

```

|||||
581 ATTACTAGGCTAGGATA 564

Db RESULT 25
AF534679
LOCUS AF534679 805 bp DNA linear PLN 02-SEP-2003
DEFINITION Pyricularia edulis chloroplast trnA-Leu (trnL) gene, partial sequence
and trnL-trnF intergenic spacer region.
ACCESSION AF534679
VERSION AF534679.1 GI:33330318
KEYWORDS chloroplast Pyricularia edulis
SOURCE Pyricularia edulis
ORGANISM Pyricularia edulis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Santalales; Santalaceae; Pyricularia.
REFERENCE 1 (bases 1 to 805)
AUTHORS Yang,S.-X., Yang,J.-B., Peng,H. and Li,D.-Z.
TITLE The systematic relationship of Pentaphragmataceae: inference from
trnL-F (cpDNA) sequence data
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 805)
AUTHORS Yang,S.-X., Yang,J.-B., Peng,H. and Li,D.-Z.
TITLE Direct Submission
JOURNAL Direct Submission
TITLE Institute of Botany, Heilongtan, Kunming, Yunnan 650204, P.R. China
FEATURES
source 1..805
misc_feature <1..>805
/organism="Pyricularia edulis"
/organelle="plastid:chloroplast"
/mol_type="genomic DNA"
/db_xref="taxon:212710"
/notes="contains trnA-Leu (trnL) gene and trnL-trnF
intergenic spacer region"

ORIGIN
Alignment Scores: 1.71e+03 Length: 805
Pred. No.: 27.00 Matches: 6
Score: 27.00 Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 8

US-10-014-101B-41 (1-6) x AF534679 (1-805)

Qy 1 IleThrArgAlaArgile 6
|||||
Db 430 ATTACTGGCGGAGAATA 447

RESULT 26
ZMU82201
LOCUS ZMU82201 816 bp mRNA linear PLN 06-JUL-1998
DEFINITION Zea mays pathogenesis related protein-5 (PR-5) mRNA, complete cds.
ACCESSION Z82201
VERSION U82201.1 GI:3290005
KEYWORDS Zea mays
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 816)
AUTHORS Morris,S.W., Vernooij,B., Titatarn,S., Starrett,M., Thomas,S.,
Wiltse,C.C., Frederiksen,R.A., Bhandhufalck,A., Hulbert,S. and
Uknes,S.
TITLE Induced resistance responses in maize
JOURNAL Mol. Plant Microbe Interact. 11 (7), 643-658 (1998)
MEDLINE 98313983
PUBMED 9650297
REFERENCE 2 (bases 1 to 816)
AUTHORS Morris,S.W., Vernooij,B., Hulbert,S., Starrett,M., Thomas,S. and
Uknes,S.
Direct Submission
Submitted (13-DEC-1996) Ciba-Geigy, P.O. Box 12257, Research
Triangle Park, NC 27709-2257, USA
FEATURES
source 1..816
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="taxon:4577"
gene 1..816
CDS 63..581
/gene="PR-5"
/gene="PR-5"
/codon_start=1
/product="pathogenesis related protein-5"
/protein_id="AAC25630.1"
/db_xref="GI:3290006"
/translation="MAAASSVLLLLLAALACMSANAATFTITNNCGFTVWPAATPVG
GTLQNPQGTWTVNPAGTSGRVWGRGTCSPNGSGCQTGDCGALACTLSGGPPL
TLAEFTIGGSQDFYDISVIDGYNLAWAFSCSTGVRLVCTDPCGPDAYHNPDMKTHAC
GGSNYSYQVTFCP"

ORIGIN
Alignment Scores: 1.74e+03 Length: 816
Pred. No.: 27.00 Matches: 6
Score: 27.00 Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 8

US-10-014-101B-41 (1-6) x ZMU82201 (1-816)

Qy 1 IleThrArgAlaArgile 6
|||||
Db 675 ATTACAGTGCAGCATT 692

RESULT 27
BV065128/c
LOCUS BV065128/c 860 bp DNA linear STS 31-MAY-2003
DEFINITION S212P60513PE11.T0 CZECHII/E1 Mus musculus STS genomic, sequence
tagged site.
ACCESSION BV065128
VERSION BV065128.1 GI:31180923
KEYWORDS STS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 860)
AUTHORS Wade,C.M., Kulbokas,E.J. III, Kirby,A.W., Zody,M.C., Mullikin,J.C.,
Lander,E.S., Lindblad-Toh,K. and Daly,M.J.
TITLE The mosaic structure of variation in the laboratory mouse genome
JOURNAL Nature 420 (6915), 574-578 (2002)
MEDLINE 22354684
PUBMED 12466852
COMMENT
Contact: Kerstin Lindblad-Toh
Whitehead Institute for Biomedical Research, Center for Genome
Research
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172521477
Fax: 6172580903
Email: kersli@genome.wi.mit.edu
Primer A: None
Primer B: None
STS size: 860
Protocol:
WGS-discovery: Paired-end low-coverage whole genome shotgun reads
were generated from 129S1/SvimJ, C3H/HeJ, and BALB/cByJ. The WGS
reads were placed uniquely on the MGSv3 C57BL/6J assembly and SNP
detection was carried out by SSAHA-SNP. 225,000 reads were
annotated

```

as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J and the strain from which the particular read came. The validation rate for these SNPs was estimated at approximately 98%.

## FEATURES

## source

```

1. .860
  /organism="Mus musculus"
  /mol_type="genomic DNA"
  /strain="CZECHII/Ei"
  /db_xref="taxon:10090"
  /map="4 2 28-808 147970603-147971383"
  /clone_lib="CZECHII/Ei"
<1. .860

```

## STS

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.83e+03 Length: 860  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 11 Gaps: 0

US-10-014-101B-41 (1-6) x BV065128 (1-860)

Qy 1 IlleThrArgAlaArglle 6

Db 43 ATCACAGGCCAGATA 26

## RESULT 28

## AK103395

LOCUS ORyza sativa (japonica cultivar-group) cDNA clone:J033127M23, full insert sequence.

## ACCESSION

VERSION AK103395.1 GI:32988604

## KEYWORDS

FLI CDNA; CAP trapper.

## SOURCE

ORyza sativa (japonica cultivar-group)

ORyza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.

## REFERENCE

## AUTHORS

The Rice Full-length cDNA Consortium, National Institute of  
 Agrobiological Sciences Rice Full-length cDNA Project Team;  
 Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,  
 Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H.,  
 Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,  
 Ohtsuki, K., Shishiki, T., Foundation of Advancement of International  
 Science Genome Sequencing & Analysis Group; Ohtsuki, K.,  
 Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,  
 Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,  
 Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,  
 Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,  
 Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN;  
 Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,  
 Hara, A., Hashizume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,  
 Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,  
 Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
 Yoshino, M., and Hayashizaki, Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

2 (bases 1 to 872)  
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,  
 Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,  
 Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,  
 Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,  
 Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,  
 Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,  
 Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,  
 Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,

Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,  
 Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,  
 Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,  
 Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K.,  
 Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,  
 Osato, N., Ota, Y., Ohtsuki, T., Ryo, R., Saitoh, H., Sakai, C., Sakai, K.,  
 Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,  
 Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,  
 Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,  
 Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,  
 Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,  
 Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and  
 Yoshimura, A.

## Direct Submission

## TITLE

## JOURNAL

Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of  
 Agrobiological Sciences, Department of Molecular Genetics, Head of  
 Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki

305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp,

Tel:81-29-838-7007, Fax:81-29-838-7007)  
 This clone is one of the 28K full-length cDNA clones from japonica

## COMMENT

rice.  
 URL : http://cdna01.dna.affrc.go.jp/cdna/

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,  
 Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,  
 Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., and  
 Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and  
 Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Ohtsuki, Y., Iida, Y.,  
 Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,  
 Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,  
 Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,  
 Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,  
 Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center  
 and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,

Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,  
 Hirooka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,  
 Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,  
 Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,

Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N.,  
 Ota, Y., Satoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,  
 Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
 Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,  
 Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,

Yasunishi, A. and Hayashizaki, Y.

## FEATURES

## Location/Qualifiers

## 1. .872

## source

```

/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J033127M23"

```

## ORIGIN

## Alignment Scores:

Pred. No.: 1.85e+03 Length: 872  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 8 Gaps: 0

US-10-014-101B-41 (1-6) x AK103395 (1-872)

Qy 1 IlleThrArgAlaArglle 6

Db 206 ATTACGCGCGCAAGATC 223

## RESULT 29

## BD275463

## LOCUS

ORyza sativa (japonica cultivar-group) cDNA clone:J033127M23, full insert sequence.

US-10-014-101B-41 (1-6) x AK103395 (1-872)

Qy 1 IlleThrArgAlaArglle 6

Db 206 ATTACGCGCGCAAGATC 223

RESULT 29

BD275463

LOCUS

```

DEFINITION 47 Human Secreted Proteins.
ACCESSION BD275463
VERSION BD275463.1 GI:33085231
KEYWORDS JP 2002539831-A/21.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 874)
AUTHORS Rosen,C.A., Ruben,S.M. and Komatsoulis,G.
TITLE 47 Human Secreted Proteins
JOURNAL Patent: JP 2002539831-A 21 26-NOV-2002;
Homo Genome Sciences Inc
COMMENT OS Homo sapiens
PN JP 2002539831-A/21
PD 26-NOV-2002
PF 22-MAR-2000 JP 2000608749
PI 60/171550,26-MAR-1999 US 60/126600 PI
craig a rosen,steven m ruben,george komatsoulis CC
FH Key Location/Qualifiers
FEATURES
    source
        1..874
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 1.85e+03 Length: 874
Score: 27.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-014-101B-41 (1-6) x BD275463 (1-874)
Qy 1 IleThrArgAlaArgile 6
Db 717 ATAACAGGGCTAGAA 734
RESULT 30
CNS06ELJ 887 bp DNA linear STS 10-JAN-2001
LOCUS T7 end of clone AROAA022E09 of library AROAA from strain CBS 732 of
DEFINITION Zygosaccharomyces rouxii, sequence tagged site.
ACCESSION AL395261
VERSION AL395261.1 GI:12146730
KEYWORDS STS.
SOURCE Zygosaccharomyces rouxii
ORGANISM Zygosaccharomyces rouxii
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Zygosaccharomyces.
REFERENCE 1 (bases 1 to 887)
AUTHORS Souciet,J.-L., Aigle,M., Artiguenave,F., Blandin,G.,
Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Illorente,B.,
Malpertuy,A., Neuvéglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
Saurin,W., Tekaiia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
Wincker,P. and Weissenbach,J.
TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of
yeast species for molecular evolution studies
JOURNAL FEBS Lett. 487 (1), 3-12 (2000)
MEDLINE 20584711
PUBMED 11152876
REFERENCE 2 (bases 1 to 887)
AUTHORS de Montigny,J., Straub,M., Potier,S., Tekaiia,F., Dujon,B.,
Wincker,P., Artiguenave,F. and Souciet,J.
TITLE Genomic exploration of the hemiascomycetous yeasts: 8.
Zygosaccharomyces rouxii
JOURNAL FEBS Lett. 487 (1), 52-55 (2000)
MEDLINE 20584718
PUBMED 11152883
REFERENCE 3 (bases 1 to 887)
AUTHORS Genoscope - Centre National de Sequencage,
Direct Submission
TITLE Submitted (06-SEP-2000) Genoscope - Centre National de Sequencage,
JOURNAL 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail:
secref@genoscope.cns.fr - Web: www.genoscope.cns.fr)
COMMENT This STS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of
the other extremity of this insert.
FEATURES
    source
        1..887
        /organism="Zygosaccharomyces rouxii"
        /mol_type="genomic DNA"
        /strain="CBS 732"
        /db_xref="taxon:4956"
        /clone="AROAA022E09"
        /clone_lib="AR0AA"
        /note="end: 17"
    misc_feature
        <3..>407
        /note="similar to Saccharomyces cerevisiae ORF YPR173c [
        VPS4 ; vacuolar sorting protein ]"
    misc_feature
        complement(<539..>886)
        /note="similar to Saccharomyces cerevisiae ORF YLR456w [
        strong similarity to YPR172w ]"
    misc_feature
        /evidence=not_experimental
        complement(<548..>886)
        /note="similar to Saccharomyces cerevisiae ORF YPR172w [
        strong similarity to YLR456w ]"
        /evidence=not_experimental
ORIGIN
Alignment Scores:
Pred. No.: 1.88e+03 Length: 887
Score: 27.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 11 Gaps: 0
US-10-014-101B-41 (1-6) x CNS06ELJ (1-887)
Qy 1 IleThrArgAlaArgile 6
Db 281 ATTACAAGAGCCAGAA 298
RESULT 31
E16279 900 bp DNA linear PAT 28-JUL-1999
LOCUS gDNA encoding ubiquitin.
DEFINITION E16279
ACCESSION E16279
VERSION E16279.1 GI:5710962
KEYWORDS JP 1998150989-A/1.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 900)
AUTHORS Tomota,A. and Koike,M.
TITLE UBIQUITIN GENE
JOURNAL Patent: JP 1998150989-A 1 09-JUN-1998;
NOYAKU BIO TECHNOL KAIHATSU GIJUTSU KENKYU KUMIAI
COMMENT OS Mamestra brassicae nuclear polyhedrosis virus PN JP
1998150989-A/1
PD 09-JUN-1998
PF 26-NOV-1996 JP 1996315069
PI TOMOTA AKIHIRO, KOIKE MASARU

```

```

PC C12N15/09.C07H21/04//C07K14/01.(C12N15/09,C12R1:92); CC
strandedness: Double;
CC topology: linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FT source 1..900
FT /organism='Mamestra brassicae nuclear FT
FT polyhedrosis virus'
FT 5'UTR 1..405
FT CDS 406..708
FT /product='Ubiquitin'
FT 3'UTR 709..900.
FT Location/Qualifiers
FEATURES
source
1..900
/organism='unidentified'
/mol_type='genomic DNA'
/db_xref='taxon:32644'

ORIGIN
Alignment Scores:
Pred. No.: 1.91e+03 Length: 900
Score: 27.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-014-101B-41 (1-6) x E16279 (1-900)

QY 1 IleThrArgAlaArgile 6
Db 656 ATAACCGGGCTCGAATA 673

RESULT 32
AF434177/c 984 bp DNA linear MAM 17-OCT-2002
LOCUS
DEFINITION
Thylamys elegans isolate NK 95622 cytochrome b (cyt b) gene,
partial cds; mitochondrial gene for mitochondrial product.
ACCESSION
AF434177
VERSION
AF434177.1 GI:24078500
KEYWORDS
mitochondrion Thylamys elegans (Elegant Fat-tailed Opossum)
SOURCE
Thylamys elegans
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Metatheria; Didelphimorphia; Didelphidae; Thylamys.
REFERENCE
1 (bases 1 to 984)
AUTHORS
Meynard,A.P., Palma,R.E. and Rivera-Milla,E.
TITLE
Flogeografia de las llacas chilenas del genero Thylamys
(Warupialia, Didelphidae) en base a secuencias del gen
mitochondrial citocromo b
JOURNAL
Rev. Chil. Hist. Nat. (2002) In press
REFERENCE
2 (bases 1 to 984)
AUTHORS
Meynard,A.P., Palma,R.E. and Rivera-Milla,E.
TITLE
Direct Submission
JOURNAL
Submitted (16-OCT-2001) Ecologia, P. Universidad Catolica de Chile,
Casilla 114-D, Santiago, Santiago 6513677, Chile
FEATURES
source
1..984
/organism='Thylamys elegans'
/organelle='mitochondrion'
/mol_type='genomic DNA'
/db_xref='taxon:191871'
<1..984
/gene='cyt b'
<1..984
/gene='cyt b'
/codon_start=1
/transl_table=2
/product='cytochrome b'
/protein_id='AA045852.1'

```

```

/db_xref='GI:24078501'
/translation='MTHLRKSHPLLKIIINHSFIDLPAISPNSISAWNFGSLGVCLLIQ
ILTGLFLAMHYTSDTAFSSVAHILIRNIHANGASMFMCFLFVGRGL
YGYSLFKFTWNVGVILLTWTATFVGVYLVPMQMSFWGATVITLLSAIPYIGTLL
VEWIMGGFSVDKATLRFPAFHFILPFIILALVIVHLLFLHETGSNNPTGINDSDKI
PHPYVYIKDVLGLIIMILMLLTALFSPDTLGDNDFTFANLTPPHIKTWYFLF
AYALLWSIPNKGVLALLASILVLLINSTLHTSNQWSLMPFISQTLFWLL"

ORIGIN
Alignment Scores:
Pred. No.: 2.08e+03 Length: 984
Score: 27.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-014-101B-41 (1-6) x AF434177 (1-984)

QY 1 IleThrArgAlaArgile 6
Db 581 ATAACAAGGGCTAGGATA 564

RESULT 33
AX935448 984 bp DNA linear PAT 05-JAN-2004
LOCUS
DEFINITION
Sequence 52 from Patent WO03089475.
ACCESSION
AX935448
VERSION
AX935448.1 GI:40642196
KEYWORDS
synthetic construct
SOURCE
synthetic construct
ORGANISM
other sequences; artificial sequences.
REFERENCE
1
AUTHORS
Peschen,D., Fischer,R., Schillberg,S., Liao,Y.C. and Dorfmueller,S.
TITLE
Antibodies, recombinant antibodies, recombinant antibody fragments
and fusions mediated plant disease resistance against fungi
JOURNAL
Patent: WO 03089475-A 52 30-OCT-2003;
Fraunhofer-Gesellschaft zur Foerderung der angewandten Forschung
e.V. (DE)
FEATURES
Location/Qualifiers
source
1..984
/organism='synthetic construct'
/mol_type='unassigned DNA'
/db_xref='taxon:32630'
/notes='Description of Artificial Sequence: precursor
fusion protein comprising the RS - linker - scFv CWP2.'

ORIGIN
Alignment Scores:
Pred. No.: 2.08e+03 Length: 984
Score: 27.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-014-101B-41 (1-6) x AX935448 (1-984)

QY 1 IleThrArgAlaArgile 6
Db 59 ATAACCGGTGCAAGATC 76

RESULT 34
AF434176/c 985 bp DNA linear MAM 17-OCT-2002
LOCUS
DEFINITION
Thylamys elegans isolate NK 96571 cytochrome b (cyt b) gene,
partial cds; mitochondrial gene for mitochondrial product.
ACCESSION
AF434176
VERSION
AF434176.1 GI:24078498
KEYWORDS
mitochondrion Thylamys elegans (Elegant Fat-tailed Opossum)
SOURCE
Thylamys elegans

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Metatheria; Didelphimorphia; Didelphidae; Thylamys.

## REFERENCE

1 (bases 1 to 985)

## AUTHORS

Meynard,A.P., Palma,R.E. and Rivera-Milla,E.

## TITLE

Filogeografía de las llacras chilenas del genero Thylamys

## JOURNAL

mitochondrial citocromo b

## REFERENCE

Rev. Chil. Hist. Nat. (2002) In press

## AUTHORS

Meynard,A.P., Palma,R.E. and Rivera-Milla,E.

## TITLE

Direct Submission

## JOURNAL

Submitted (16-Oct-2001) Ecologia, P. Universidad Catolica de Chile,

## FEATURES

Casilla 114-D, Santiago, Santiago 6513677, Chile

## source

Location/Qualifiers

## 1. .985

/organism="Thylamys elegans"

/organelle="mitochondrion"

/mol\_type="genomic DNA"

/isolate="NK 96571"

/db\_xref="taxon:191871"

## gene

<1..>985

## CDS

/genes="cyt b"

<1..>985

/genes="cyt b"

/codon\_start=1

/transl\_table=2

/product="cytochrome b"

/protein\_id="AA045851.1"

/db\_xref="GI:24078499"

/translation="WTHLRKSHPLKLNHSDIDLPAPSNISAWNFGSLGVLIIQ

ITLGLFLAHYSDTSFSSVAHICRDVNFGLNRNIHANGASFFMCLFLHVGRL

YTGSLFKEFTWNGVILLTVMATFVGVLPGQSPFWGATVITNLSAIPYIGTL

VEWIMGSGVDKATRRFAFHILFPFIILALIVHLLFHTGNNPSGPNPDSDKI

PHPPYTIKDVGLIIMLLMLTLALFSDTLGDPNFTPANPLNTPPHIKPEWYFLF

AVALLWSIPNKGVLALLASILVLLASSTTSYIKETKLINSTNFTNLVLI"

## ORIGIN

## Alignment Scores:

Pred. No.: 2.08e+03 Length: 985  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-41 (1-6) x AF434176 (1-985)

## QY

1 IleThrArgAlaArgIle 6

## Db

581 ATAAACAGCGCTAGGATA 564

## RESULT 35

AX935453

LOCUS AX935453

DEFINITION Sequence 57 from Patent WO03089475.

ACCESSION AX935453

VERSION AX935453.1 GI:40642201

KEYWORDS

SOURCE

ORGANISM

1

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. .987

/organism="synthetic construct"

/mol\_type="unassigned DNA"

/db\_xref="taxon:32630"

## ORIGIN

## Alignment Scores:

Pred. No.: 2.09e+03 Length: 987  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-014-101B-41 (1-6) x AX935453 (1-987)

## QY

1 IleThrArgAlaArgIle 6

## Db

59 ATAAACGCGTGCAGGAATC 76

## RESULT 36

AX935455

LOCUS AX935455

DEFINITION Sequence 59 from Patent WO03089475.

ACCESSION AX935455

VERSION AX935455.1 GI:40642203

KEYWORDS

SOURCE

ORGANISM

1

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. .987

/organism="synthetic construct"

/mol\_type="unassigned DNA"

/db\_xref="taxon:32630"

/note="Description of Artificial Sequence: precursor fusion protein comprising RS - linker scFv VDM2."

## ORIGIN

## Alignment Scores:

Pred. No.: 2.09e+03 Length: 987  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-014-101B-41 (1-6) x AX935455 (1-987)

## QY

1 IleThrArgAlaArgIle 6

## Db

59 ATAAACGCGTGCAGGAATC 76

## RESULT 37

AF431925/c

LOCUS AF431925/c

DEFINITION Thylamys elegans isolate NK 95436 cytochrome b (cytb) gene, partial cds; mitochondrial gene for mitochondrial product.

ACCESSION AF431925

VERSION AF431925.1 GI:24849957

KEYWORDS

SOURCE

ORGANISM

1

REFERENCE

AUTHORS

1 (bases 1 to 988)

Eduardo Palma,R., Rivera-Milla,E., Yates,T.L., Marquet,P.A. and

Meynard,A.P.

/note="Description of Artificial Sequence: precursor fusion protein comprising RS - linker - scFv VDCw."

## Alignment Scores:

Pred. No.: 2.09e+03 Length: 987  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-014-101B-41 (1-6) x AX935453 (1-987)

## QY

1 IleThrArgAlaArgIle 6

## Db

59 ATAAACGCGTGCAGGAATC 76

## RESULT 36

AX935455

LOCUS AX935455

DEFINITION Sequence 59 from Patent WO03089475.

ACCESSION AX935455

VERSION AX935455.1 GI:40642203

KEYWORDS

SOURCE

ORGANISM

1

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

Location/Qualifiers

1. .987

/organism="synthetic construct"

/mol\_type="unassigned DNA"

/db\_xref="taxon:32630"

/note="Description of Artificial Sequence: precursor fusion protein comprising RS - linker scFv VDM2."

## ORIGIN

## Alignment Scores:

Pred. No.: 2.09e+03 Length: 987  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-014-101B-41 (1-6) x AX935455 (1-987)

## QY

1 IleThrArgAlaArgIle 6

## Db

59 ATAAACGCGTGCAGGAATC 76

## RESULT 37

AF431925/c

LOCUS AF431925/c

DEFINITION Thylamys elegans isolate NK 95436 cytochrome b (cytb) gene, partial cds; mitochondrial gene for mitochondrial product.

ACCESSION AF431925

VERSION AF431925.1 GI:24849957

KEYWORDS

SOURCE

ORGANISM

1

REFERENCE

AUTHORS

1 (bases 1 to 988)

Eduardo Palma,R., Rivera-Milla,E., Yates,T.L., Marquet,P.A. and

Meynard,A.P.

**TITLE** Phylogenetic and biogeographic relationships of the mouse opossum  
**JOURNAL** Thylamys (Didelphimorphia, Didelphidae) in southern South America  
**MEDLINE** Mol. Phylogenet. Evol. 25 (2), 245-253 (2002)  
**PUBLISHED** 22302519  
**REFERENCE** 2 (bases 1 to 988)  
**AUTHORS** Palma, R. Eduardo., Rivera-Milla, E., Yates, T.L., Marquet, P.A. and  
Meynard, A.P.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (11-OCT-2001) Ecologia, P. Universidad Catolica de Chile,  
Casilla 114-D, Santiago 6513677, Chile  
**FEATURES** Location/Qualifiers  
source  
1..988  
/organism="Thylamys elegans"  
/organelle="mitochondrion"  
/mol\_type="genomic DNA"  
/isolate="NK 95436"  
/db\_xref="taxon:191871"  
<1..>988  
/genes="cytb"  
<1..>988  
/genes="cytb"  
/codon\_start=1  
/transl\_table=2  
/product="cytochrome b"  
/protein\_id="AAN64835.1"  
/db\_xref="GI:24849958"  
/translation="MTLARKSHPLKIIINHSPDLPAPSNISAWNFGSLGVLCLIIQ  
ILTGLFLAMHYTSDTLTAFSSVAHICRDVNFGLIRNIHANGASFMFCLFLHVGRL  
YVGSYLFKFTWNVGILLTVMTAFVGVLPWQMSFWGATVITNLSAIPYIGTTL  
VEWINGGFSVDKATLTFRFAFHILPFIILALVIVHLLFHTGNNPTGPNDSDKI  
PFPYTTIKDVLGLIIMLLLTALFSPDMLGDPNFTPLNPATPPHKEWYFLF  
AVAILRSIPNKGVLALLASTLVLLVIVIPLLHTSNQRLMFRPISQTLFWLLTA"  
**gene**  
**CDS**  
**Alignment Scores:**  
Pred. No.: 2.09e+03 Length: 988  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0  
**US-10-014-101B-41 (1-6) x AF431925 (1-988)**  
**QY** 1 IleThrArgAlaArgIle 6  
|||||  
**Db** 581 ATACACAGGCTAGGATA 564  
**RESULT 38**  
**LOCUS** AF431923/c 990 bp DNA linear MAM 10-NOV-2002  
**DEFINITION** Thylamys pallidior isolate EP 440 cytochrome b (cytb) gene, partial  
cds; mitochondrial gene for mitochondrial product.  
**ACCESSION** AF431923  
**VERSION** AF431923.1 GI:24849953  
**KEYWORDS**  
**SOURCE** mitochondrion Thylamys pallidior (pallid fat-tailed opossum)  
**ORGANISM** Thylamys pallidior  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Metatheria; Didelphimorphia; Didelphidae; Thylamys.  
**REFERENCE** 1 (bases 1 to 990)  
**AUTHORS** Eduardo Palma, R., Rivera-Milla, E., Yates, T.L., Marquet, P.A. and  
Meynard, A.P.  
**TITLE** Phylogenetic and biogeographic relationships of the mouse opossum  
**JOURNAL** Thylamys (Didelphimorphia, Didelphidae) in southern South America  
**MEDLINE** Mol. Phylogenet. Evol. 25 (2), 245-253 (2002)  
**PUBLISHED** 22302519  
**REFERENCE** 2 (bases 1 to 990)  
**AUTHORS** Palma, R. Eduardo., Rivera-Milla, E., Yates, T.L., Marquet, P.A. and  
Meynard, A.P.  
**TITLE** Direct Submission

**JOURNAL** Submitted (11-OCT-2001) Ecologia, P. Universidad Catolica de Chile,  
Casilla 114-D, Santiago 6513677, Chile  
**FEATURES** Location/Qualifiers  
source  
1..990  
/organism="Thylamys pallidior"  
/organelle="mitochondrion"  
/mol\_type="genomic DNA"  
/isolate="EP 440"  
/db\_xref="taxon:126301"  
<1..>990  
/genes="cytb"  
<1..>990  
/genes="cytb"  
/codon\_start=1  
/transl\_table=2  
/product="cytochrome b"  
/protein\_id="AAN64833.1"  
/db\_xref="GI:24849954"  
/translation="MTNLRKSHPLKIIINHSPDLPAPSNISAWNFGSLGVLCLIIQ  
ILTGLFLAMHYTSDTLTAFSSVAHICRDVNFGLIRNIHANGASFMFCLFLHVGRL  
YVGSYLFKFTWNVGILLTVMTAFVGVLPWQMSFWGATVITNLSAIPYIGTTL  
VEWINGGFSVDKATLTFRFAFHILPFIILALVIVHLLFHTGNNPTGPNDSDKI  
PFPYTTIKDVLGLIIMLLLTALFSPDMLGDPNFTPLNPATPPHKEWYFLF  
AVAILRSIPNKGVLALLASTLVLLVIVIPLLHTSNQRLMFRPISQTLFWLLTA"  
**ORIGIN**  
**Alignment Scores:**  
Pred. No.: 2.09e+03 Length: 990  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0  
**US-10-014-101B-41 (1-6) x AF431923 (1-990)**  
**QY** 1 IleThrArgAlaArgIle 6  
|||||  
**Db** 581 ATACACAGGCTAGGATA 564  
**RESULT 39**  
**LOCUS** CQ497388 1039 bp DNA linear PAT 30-JAN-2004  
**DEFINITION** Sequence 29255 from Patent WO0160860.  
**ACCESSION** CQ497388  
**VERSION** CQ497388.1 GI:41463024  
**KEYWORDS**  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
**REFERENCE** 1  
**AUTHORS** Schlegel, R., Endege, W.O. and Monahan, J.E.  
**TITLE** Genes differentially expressed in human prostate cancer and their  
use  
Patent: WO 0160860-A 29255 23-AUG-2001;  
Millennium Predictive Medicine, Inc. (US)  
**JOURNAL** Location/Qualifiers  
**FEATURES** source  
1..1039  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
**ORIGIN**  
**Alignment Scores:**  
Pred. No.: 2.19e+03 Length: 1039  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
**US-10-014-101B-41 (1-6) x CQ497388 (1-1039)**

QY 1 IleThrArgAlaArgile 6  
|||||  
Db 892 ATTACGAGGCGACGAATA 909

RESULT 40  
AY333784 1045 bp DNA linear INV 07-JUN-2004  
LOCUS Toxoplasma gondii satellite TgSat350-4 sequence.  
DEFINITION AY333784  
ACCESSION AY333784  
VERSION AY333784.1 GI:33188107  
KEYWORDS  
SOURCE Toxoplasma gondii  
ORGANISM Toxoplasma gondii  
Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;  
Sarcocystidae; Toxoplasma.  
REFERENCE 1 (bases 1 to 1045)  
AUTHORS Clemente, M., De Miguel, N., Lia, V. V., Matrajt, M. and Angel, S. O.  
TITLE Structure Analysis of Two Toxoplasma gondii and Neospora caninum  
Satellite DNA Families and Evolution of Their Common Monomeric  
Sequence  
JOURNAL J. Mol. Evol. 58 (5), 557-567 (2004)  
PUBMED 15170259  
REFERENCE 2 (bases 1 to 1045)  
AUTHORS Clemente, M., de Miguel, N. and Angel, S. O.  
TITLE Direct Submission  
JOURNAL Submitted (01-JUL-2003) Parasitologia Molecular, IIB-INTECH, Camino  
de Circunvalacion Laguna Km 6, Chascomus, Buenos Aires B7130IWA,  
Argentina

FEATURES  
source 1..1045  
/organism="Toxoplasma gondii"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:5811"  
repeat\_region 1..1045  
/note="satellite TgSat 350-4"  
/rpt\_type=tandem

ORIGIN  
Alignment Scores:  
Pred. No.: 2.2e+03 Length: 1045  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0  
US-10-014-101b-41 (1-6) x AY333784 (1-1045)

QY 1 IleThrArgAlaArgile 6  
|||||  
Db 773 ATCACTCGGCGACGAATT 790

Search completed: February 18, 2005, 05:26:06  
Job time : 828.72 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 17, 2005, 21:47:40 ; Search time 107.16 Seconds  
(without alignments)  
331.452 Million cell updates/sec

Title: US-10-014-101b-41  
Perfect score: 27  
Sequence: 1 ITRARI 6

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2 1/USPTO spo04 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0  
-DB=N Geneseq 16Dec04 -START=1 -END=1 -MATRIX=Blosum62 -TRANS=human40.cdi  
-LOPEXT=0 -UNITS=bits  
-LIST=45 -DOCALIGN=200 -THR SCORE=spct -THR MAX=100 -THR MIN=0 -ALIGN=40  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10014101 @CGN 1 1 701 @runat 16022005 075845 7925 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	100.0	159	10	ADG32304
c	2	27	100.0	192	6 ABN68746
c	3	27	100.0	195	6 ABN70489
4	27	100.0	354	10	ACD94249
5	27	100.0	375	4	ABA09353

c	6	27	100.0	384	8	ABX53359	Abx53359 Bovine ES
	7	27	100.0	493	5	ABV07966	Abv07966 Human pro
	8	27	100.0	503	3	AAC36788	Aac36788 Arabidops
	9	27	100.0	537	6	ABQ24115	Abq24115 Oligonuel
c	10	27	100.0	537	6	ABQ24114	Abq24114 Oligonuel
	11	27	100.0	708	8	ACA21667	Aca21667 Prokaryot
	12	27	100.0	831	2	AAQ99805	Aaq99805 Thaumatin
	13	27	100.0	874	3	AAC69410	Aac69410 Human sec
c	14	27	100.0	900	2	AAV32472	Aav32472 Ubiquitin
	15	27	100.0	936	11	ABD13501	Abd13501 Pseudomon
	16	27	100.0	984	10	ADG32343	Adg32343 DNA encod
	17	27	100.0	987	10	ADG32350	Adg32350 DNA encod
	18	27	100.0	987	10	ADG32348	Adg32348 DNA encod
	19	27	100.0	1039	5	ABV29237	Abv29237 Human pro
	20	27	100.0	1059	10	ADC24030	Adc24030 DNA seque
	21	27	100.0	1059	12	ADH36131	Adh36131 Chemical
	22	27	100.0	1059	12	ADG93832	Adg93832 Nitrlase
	23	27	100.0	1059	12	ADI62429	Adi62429 DNA encod
	24	27	100.0	1059	12	ADI64550	Adi64550 DNA encod
c	25	27	100.0	1124	13	ADS60562	Ads60562 Bacterial
c	26	27	100.0	1171	8	ACA21136	Aca21136 Prokaryot
c	27	27	100.0	1215	8	ACA01962	Aca01962 C. glutam
	28	27	100.0	1290	11	ABD13808	Abd13808 Pseudomon
c	29	27	100.0	1311	11	ABD13579	Abd13579 Pseudomon
c	30	27	100.0	1434	5	AAH65248	Aah65248 C. glutam
	31	27	100.0	1436	3	AAC44665	Aac44665 Arabidops
	32	27	100.0	1506	6	ABK28625	Abk28625 cDNA enco
	33	27	100.0	1506	9	ACC85294	Acc85294 Arabidops
	34	27	100.0	1506	10	ADH61267	Adh61267 Arabidops
	35	27	100.0	1506	12	ADO06498	Ado06498 A thalian
	36	27	100.0	1515	6	ABK28629	Abk28629 cDNA enco
	37	27	100.0	1515	9	ACC85298	Acc85298 Arabidops
	38	27	100.0	1539	8	ACA26612	Aca26612 Prokaryot
	39	27	100.0	1548	3	AAC43214	Aac43214 Arabidops
c	40	27	100.0	1564	11	ADL66030	Adl66030 C. glutam
	41	27	100.0	1566	8	ADA71115	Ada71115 Rice gene
	42	27	100.0	1572	6	ABK28626	Abk28626 cDNA enco
	43	27	100.0	1572	9	ACC85295	Acc85295 Arabidops
	44	27	100.0	1572	10	ADH61268	Adh61268 Arabidops
	45	27	100.0	1575	3	AAC42983	Aac42983 Arabidops

ALIGNMENTS

RESULT 1  
ADG32304  
ID ADG32304 standard; cDNA; 159 BP.  
XX  
AC ADG32304;  
XX  
DT 26-FEB-2004 (first entry)  
XX  
DE Radish cDNA encoding an antifungal protein (AFP) SeqID 13.  
XX  
KW radish; scfV; gene; ss; anti-fungal peptide; AFP; scfV;  
XX disease resistant; transgenic; plant; fungal infection; antibody;  
XX pathogen-specific antibody; fungicidal; agriculture.  
OS Raphanus sativus.  
XX  
FN WO2003089475-A2.  
XX  
PD 30-OCT-2003.  
XX  
PF 14-APR-2003; 2003WO-EP003852.  
XX  
PR 22-APR-2002; 2002EP-00008929.  
PR 28-MAY-2002; 2002EP-00011807.  
XX  
(FRAU ) FRAUNHOFER GES FOERDERUNG ANGEWANDTEN.  
PI Peschen D, Fischer R, Schillberg S, Liao Y, Dorfmueller S;  
XX

DR WPI; 2003-854088/79.  
 DR P-PSDB; ADG32329.  
 XX  
 PT New fusion protein comprising an anti-fungal protein or peptide and an  
 PT antibody fragment, useful in agriculture and horticulture for producing  
 PT Ascomyceta-resistant transgenic plants, plant cells or plant tissues.  
 XX  
 XX  
 PS Claim 26; SEQ ID NO 13; 47pp; English.  
 XX  
 CC This invention relates to a novel fusion protein comprising an anti-  
 CC fungal protein or peptide (APP) and an antibody fragment (scFv).  
 CC Specifically, it refers to APPs recognizing an epitope of an Ascomyceta  
 CC and a cellular targeting sequence, which can be used to generate disease  
 CC resistant transgenic plants that are protected against fungal infection.  
 CC Accordingly, a method is described for antibody based resistance in  
 CC plants such that the undesirable and expensive chemical controls often  
 CC used in agriculture are not required. The present invention provides  
 CC antibodies, recombinant antibodies and fragments thereof, as well as  
 CC fusion proteins that can be used as pathogen-specific antibodies targeted  
 CC to different plant cell compartments. As such, these fungicidal agents  
 CC confer a broad spectrum of disease resistance in both economically  
 CC important crops and ornamental plants. This polynucleotide is a cDNA  
 CC sequence encoding an antifungal protein. This polynucleotide is the invention.  
 XX  
 SQ Sequence 159 BP; 42 A; 32 C; 42 G; 43 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 526 Length: 159  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 10 Gaps: 0

US-10-014-101B-41 (1-6) x ADG32304 (1-159)

Qy 1 IleThrArgAlaArgIle 6  
 |||||  
 Db 59 ATACGGGTGCAGATC 76

RESULT 2  
 ABN68746/c  
 ID ABN68746 standard; DNA; 192 BP.

XX AC ABN68746;

XX DT 01-JUL-2002 (first entry)

XX DE Streptococcus polynucleotide SEQ ID NO 5405.

XX DE Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;  
 XX group A streptococcus; Streptococcus pyogenes; antibacterial; gene;  
 XX antiinflammatory; infection; vaccine; meningitis; gene therapy; ds.  
 XX OS Streptococcus agalactiae.

XX PN WO200234771-A2.

XX PD 02-MAY-2002.

XX PF 29-OCT-2001; 2001WO-GB004789.

XX PR 27-OCT-2000; 2000GB-00026333.

XX PR 24-NOV-2000; 2000GB-00028727.

XX PR 07-MAR-2001; 2001GB-00005640.

XX PA (CHIR-) CHIRON SPA.

XX PA (GENO-) INST GENOMIC RES.

XX PI Telford J, Massignani V, Margarit Y Rosi, Grandi G, Fraser C;

XX PI Tettelin H;

XX PI WPI; 2002-352536/38.

DR P-PSDB; ABP28115.

XX  
 PT New Streptococcus protein for the treatment or prevention of infection or  
 PT disease caused by Streptococcus bacteria, such as meningitis, and for  
 PT detecting a compound that binds to the protein.

PS Claim 7; Page 3702; 4525pp; English.

XX  
 CC The invention relates to a protein (ABP25413-ABP30895) from group B  
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
 CC the specification. The proteins have antibacterial and antiinflammatory  
 CC activity. (I), nucleic acids encoding (I), ABM6044-ABN71526 and  
 CC antibodies that bind (I) are used in the manufacture of medicaments for  
 CC the treatment or prevention of infection or disease caused by  
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
 CC biological sample. (I) is used to determine whether a compound binds to  
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be  
 CC used as a vaccine or diagnostic composition. The disease caused by  
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
 CC acid encoding (I) may be used to recombinantly produce (I) and may be  
 CC used in gene therapy. Antibodies to (I) are used for affinity  
 CC chromatography, immunoassays, and distinguishing/identifying  
 CC Streptococcus proteins

XX SQ Sequence 192 BP; 56 A; 29 C; 42 G; 65 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 647 Length: 192  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-10-014-101B-41 (1-6) x ABN68746 (1-192)

Qy 1 IleThrArgAlaArgIle 6  
 |||||  
 Db 116 ATACAGAGCAAGATT 99

RESULT 3

ABN70489/c

ID ABN70489 standard; DNA; 195 BP.

XX AC ABN70489;

XX DT 01-JUL-2002 (first entry)

XX DE Streptococcus polynucleotide SEQ ID NO 8891.

XX DE Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;  
 XX group A streptococcus; Streptococcus pyogenes; antibacterial; gene;  
 XX antiinflammatory; infection; vaccine; meningitis; gene therapy; ds.

XX OS Streptococcus agalactiae.

XX PN WO200234771-A2.

XX PD 02-MAY-2002.

XX PF 29-OCT-2001; 2001WO-GB004789.

XX PR 27-OCT-2000; 2000GB-00026333.

XX PR 24-NOV-2000; 2000GB-00028727.

XX PR 07-MAR-2001; 2001GB-00005640.

XX PA (CHIR-) CHIRON SPA.

XX PA (GENO-) INST GENOMIC RES.

XX PI Telford J, Massignani V, Margarit Y Rosi, Grandi G, Fraser C;

XX PI Tettelin H;

XX WPI; 2002-352536/38.  
 DR P-PSDB; ABP29858.  
 XX  
 PT New Streptococcus protein for the treatment or prevention of infection or  
 PT disease caused by Streptococcus bacteria, such as meningitis, and for  
 PT detecting a compound that binds to the protein.  
 XX  
 PS Claim 7; Page 4007; 4525pp; English.  
 XX  
 CC The invention relates to a protein (ABP25413-ABP30895) from group B  
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
 CC the specification. The proteins have antibacterial and antiinflammatory  
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and  
 CC antibodies that bind (I) are used in the manufacture of medicaments for  
 CC the treatment or prevention of infection or disease caused by  
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
 CC biological sample. (I) is used to determine whether a compound binds to  
 CC used as a vaccine or diagnostic composition. The disease caused by  
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
 CC acid encoding (I) may be used to recombinantly produce (I) and may be  
 CC used in gene therapy. Antibodies to (I) are used for affinity  
 CC chromatography, immunoassays, and distinguishing/identifying  
 CC Streptococcus proteins  
 XX  
 SQ Sequence 195 BP; 58 A; 29 C; 42 G; 66 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 658 Length: 195  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0  
 US-10-014-101B-41 (1-6) x ABN70489 (1-195)  
 QY 1 IleThrArgAlaArgIle 6  
 DB 116 ATACACAGACAGAGATT 99  
 RESULT 4  
 ACD94249  
 ID ACD94249 standard; cDNA; 354 BP.  
 AC ACD94249;  
 XX  
 DT 23-SEP-2003 (first entry)  
 XX  
 DE Human colon cancer cell expressed cDNA #2661.  
 KW Open reading frame detection; genome sequencing; colon cancer;  
 KW breast cancer; population genome analysis; genetic shift; cancer;  
 KW antibiotic resistance; antibiotic non-tolerance; congenital disease;  
 KW agriculture; food crop genome; resistance gene; retrovirus;  
 KW influenza virus; eukaryotic pathogen detection; trypanosome; Plasmodium;  
 KW gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2002155438-A1.  
 XX  
 PD 24-OCT-2002.  
 XX  
 PF 27-SEP-1999; 99US-00406117.  
 XX  
 PR 20-NOV-1998; 98US-00196716.  
 XX  
 PA (SIMP/) SIMPSON A J G.  
 PA (NETO/) NETO E D.

(BREN/) BRENTANI R R.  
 Simpson AJG, Neto ED, Brentani RR;  
 WPI; 2003-182626/18.  
 DR  
 XX Determining open reading frames of genome of an organism e.g. a human  
 PT suffering from cancer involves use of single oligonucleotide primer at  
 PT low stringency for preparing single-stranded cDNA from mRNA of  
 PT individual.  
 XX  
 PS Example 9; Page 400; 959pp; English.  
 XX  
 CC The invention describes a method of determining open reading frames in  
 CC the genome of organism, comprising contacting mRNA from cell of organism  
 CC with a single oligonucleotide primer (I) at low stringency, amplifying  
 CC single-stranded cDNA by reverse transcribing mRNA with (I), preparing  
 CC cDNA, sequencing the product, and repeating the contacting, preparing  
 CC and amplifying steps with different primers and sequencing resulting  
 CC nucleic acids. The method is useful for: determining that a known  
 CC nucleotide sequence from a genome of an organism corresponds to a  
 CC nucleotide sequence of an open reading frame; for preparing a contig,  
 CC nucleic acid molecule from a genome of an organism; and for sequencing  
 CC all or part of a genome of an organism. mRNA is obtained from mammalian  
 CC or human cell which is associated with a pathological condition e.g. a  
 CC colon cancer or breast cancer cell. The method is useful for analyses of  
 CC populations of subjects and can be used to carry out genetic analyses of  
 CC large or small populations. Further, it can be used to study living  
 CC systems to determine if, e.g. there have been genetic shifts which render  
 CC an individual or population more or less likely to be afflicted with  
 CC diseases such as cancer, to determine antibiotic resistance or non-  
 CC tolerance, and so forth. The method can also be used in the study of  
 CC congenital diseases, and the risk of affliction to a foetus, as well as  
 CC the study of whether the conditions are likely to be passed to offspring  
 CC through ova or sperm. The analyses for pathological conditions can be  
 CC carried out in all animals, plants, birds, fish, etc. Using this method,  
 CC in the area of agriculture, for example the genomes of food crops can be  
 CC studied to determine if resistance genes are present, defects in plant  
 CC genomes can also be studied in this way. Similarly, the method permits  
 CC determination of the pathogens which integrate into the genome, such as  
 CC retroviruses and other integrating viruses such as influenza virus, have  
 CC undergone shifts or mutations, which may require different approaches to  
 CC therapy. This method is also applied to eukaryotic pathogens, such as  
 CC trypanosomes, different types of Plasmodium, etc. The method essentially  
 CC eliminates sequencing of non-coding portions. This sequence represents a  
 CC polynucleotide isolated from human colon cancer cell cDNA library  
 XX  
 SQ Sequence 354 BP; 66 A; 100 C; 113 G; 74 T; 0 U; 1 Other;  
 Alignment Scores:  
 Pred. No.: 1.27e+03 Length: 354  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 10 Gaps: 0  
 US-10-014-101B-41 (1-6) x ACD94249 (1-354)  
 QY 1 IleThrArgAlaArgIle 6  
 DB 277 ATTACCCGAGCGGAATT 294  
 RESULT 5  
 ABA09353  
 ID ABA09353 standard; cDNA; 375 BP.  
 XX  
 AC ABA09353;  
 XX  
 DT 11-JAN-2002 (first entry)  
 XX  
 DE Human secreted protein homologue-encoding cDNA, SEQ ID NO:1129.  
 XX

KW Human; cytokine; cell proliferation; cell differentiation; growth factor;  
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;  
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;  
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;  
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;  
KW chronic inflammatory condition; proliferative retinopathy;  
KW atherosclerosis; coronary heart disease; arterial ischaemia;  
KW bone disorder; osteoporosis; vascular growth disorder;  
KW tissue regeneration; wound healing; infection; immune disorder;  
KW cell culture; drug screening; gene therapy; antiinflammatory;  
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;  
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;  
KW antifungal; vulnerary; antiulcer; ss.  
XX Homo sapiens.  
XX WO200157188-A2.  
XX 09-AUG-2001.  
XX 05-FEB-2001; 2001WO-US003800.  
XX 03-FEB-2000; 2000US-00496914.  
XX 27-APR-2000; 2000US-00560875.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Drmanac RT;  
XX WPI; 2001-457740/49.  
XX P-PSDB; ABB12109.  
XX Human proteins and DNA encoding sequences useful for preventing, treating  
XX or ameliorating a medical condition in a mammalian subject e.g. arthritis  
XX and cancer.  
XX Claim 1; Page 916; 1963pp; English.  
XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and  
XX sequences ABA08225-ABA09574 represent nucleic acids encoding them. The  
XX invention also relates to vectors and recombinant host cells comprising a  
XX nucleotide of the invention, methods of producing the novel polypeptides,  
XX antibodies against the polypeptides, methods of detecting the nucleotides  
XX or polypeptides in a sample, and methods of identifying compounds which  
XX bind to polypeptides of the invention. Although novel, many of the  
XX polypeptides of the invention have homology to known proteins, thereby  
XX giving an insight into their probable biological activities, and hence  
XX potential therapeutic applications. The polypeptides of the invention may  
XX have various activities, including cytokine, cell proliferation or cell  
XX differentiation activities; stem cell growth factor activity;  
XX haematopoiesis regulatory activity; tissue growth activity;  
XX immunomodulatory activity; activin- or inhibin-related activities;  
XX chemotactic or chemokinetic activities; haemostatic, thrombotic or  
XX thrombolytic activities; receptor or ligand activities; or may be  
XX involved in oncogenesis, cancer cell proliferation or metastasis.  
XX Depending on their biological activities, polypeptides and nucleotides of  
XX the invention are useful for preventing, treating or ameliorating medical  
XX conditions, e.g. by protein or gene therapy. Such conditions include  
XX cancers, haematopoietic disorders (e.g. myeloid or lymphoid cell  
XX disorders), chronic inflammatory conditions (e.g., asthma or arthritis),  
XX proliferative retinopathy, atherosclerosis, coronary heart disease,  
XX arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal  
XX vascular growth. Polypeptides involved with tissue regeneration and  
XX repair (or nucleic acids encoding them) may be used to promote wound  
XX healing (e.g. of burns, incisions and ulcers), while those with  
XX immunomodulatory activities may be used in the treatment of viral,  
XX bacterial and fungal infections in addition to immune disorders.  
XX Polypeptides with growth factor activity may be used in cell cultures to  
XX promote cell growth. For example, such polypeptides may be used to  
XX manipulate stem cells in culture to give rise to neuroepithelial cells  
XX that can be used to augment or replace cells damaged by illness,  
XX autoimmune disease or accidental damage. The polypeptides and nucleotides  
XX may also be used in the diagnosis of the above conditions, and in drug

CC screening techniques. The present sequence represents a cDNA encoding a  
CC novel human polypeptide of the invention  
XX  
SQ Sequence 375 BP; 73 A; 133 C; 101 G; 68 T; 0 U; 0 Other;  
Alignment Scores: Length: 375  
Pred. No.: 1.35e+03 Matches: 6  
Score: 27.00 Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 100.00% Gaps: 0  
DB: 4  
US-10-014-101B-41 (1-6) x ABA09353 (1-375)  
QY 1 lleThrArgAlaArgIle 6  
DB 224 ATCAGAGAGCTGAGATA 241  
RESULT 6  
ABX53359/c  
ID ABX53359 standard; cDNA; 384 BP.  
XX  
AC ABX53359;  
XX  
DT 25-FEB-2003 (first entry)  
XX  
DE Bovine EST associated with lactation/muscle/fat deposition #3288.  
XX  
KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;  
KW muscle deposition; fat deposition; genome mapping; gene identification;  
KW gene analysis; cattle breeding.  
XX  
OS Bos Taurus.  
XX  
PN US2002137160-A1.  
XX  
PD 26-SEP-2002.  
XX  
PF 26-OCT-2001; 2001US-00983965.  
XX  
PR 17-DEC-1998; 98US-0113678P.  
PR 15-DEC-1999; 99US-00465231.  
XX  
PA (BYAT/) BYATT J C.  
PA (MATH/) MATHIALAGAN N.  
PA (TAON/) TAO N.  
PA (WARR/) WARREN W C.  
XX  
PI Byatt JC, Mathialagan N, Tao N, Warren WC;  
XX WPI; 2003-102386/09.  
XX  
PT Purified nucleic acid molecules, useful for genome mapping, gene  
XX identification and analysis, cattle breeding or preparation of constructs  
XX for cattle gene expression and genetically improved cattle.  
XX  
PS Claim 2; SEQ ID NO 3288; 38pp; English.  
XX  
CC The invention relates to a purified nucleic acid molecule associated with  
XX lactation or muscle and fat deposition (designated LMFD), derived from  
XX cattle, and the LMFD nucleic acid can specifically hybridise to a second  
XX nucleic acid molecule comprising any of 5912 nucleotide sequences,  
XX appearing as ABX50072-ABX55983, or complements of them. Also included are  
XX ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic  
XX acid linked to a promoter and a 3' non-translated sequence that  
XX functions in the cell to cause termination of transcription and addition  
XX of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and  
XX (2) determining a level or pattern of a molecule in a bovine cell or  
XX tissue comprising: (a) incubating a marker nucleic acid (comprising any  
XX of the 5912 nucleic acid sequences or its complement or fragment) with a  
XX complementary nucleic acid molecule obtained from the bovine cell or  
XX tissue, where hybridisation between the marker nucleic acid and the

CC complementary nucleic acid permits the detection of the molecule; and (b)  
 CC detecting the level or pattern of the complementary nucleic acid, where  
 CC the detection of the complementary nucleic acid is predictive of the  
 CC level or pattern of the molecule. The LMFD nucleic acid is used for  
 CC determining a level or pattern of a molecule in a bovine cell or tissue.  
 CC It is useful for genome mapping, gene identification and analysis, cattle  
 CC breeding, preparation of constructs for use in cattle gene expression, or  
 CC for genetically improving cattle. The present sequence is one of the 5912  
 CC bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The present  
 CC sequence was not shown in the specification but was obtained in  
 CC electronic format from the USPTO web site:  
 CC seqdata.uspto.gov/sequence.html?DocID=20020137160

XX SQ Sequence 384 BP; 123 A; 58 C; 75 G; 128 T; 0 U; 0 Other;

Alignment Scores: Pred. No.: 1.38e+03 Length: 384  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 8 Gaps: 0

US-10-014-101B-41 (1-6) x ABX53359 (1-384)

QY 1 IleThrArgAlaArglle 6  
 DB 330 ATAACACAGAGCTCGGATA 313

RESULT 7

ABV07966  
 ID ABV07966 standard; cDNA; 493 BP.

XX AC ABV07966;

XX DT 13-SEP-2002 (first entry)

XX DE Human prostate expression marker cDNA 7957.

XX DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

KW pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX PN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US005171.

XX PR 17-FEB-2000; 2000US-0183319P.

XX PR 16-MAR-2000; 2000US-0189862P.

XX PR 25-MAY-2000; 2000US-0207454P.

XX PR 09-JUN-2000; 2000US-0211314P.

XX PR 18-JUL-2000; 2000US-0219007P.

XX PR 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JE;

XX DR WFI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of  
 PT prostate cells and correlating with presence of prostate cancer, useful  
 PT for detecting presence of prostate cancer, stage of prostate cancer.

XX PS Claim 1; Page 1275; 11750pp; English.

XX CC The invention relates to an isolated nucleic acid molecule (I) comprising  
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
 CC specification or its complement. (I) is useful for: (a) assessing whether  
 CC a patient is afflicted with prostate cancer; (b) monitoring the

CC progression of prostate cancer in a patient; (c) assessing the efficacy  
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
 CC determining whether prostate cancer has metastasized in a patient; (h)  
 CC assessing the aggressiveness or indolence of prostate cancer in a patient  
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

XX SQ Sequence 493 BP; 125 A; 104 C; 77 G; 132 T; 0 U; 55 Other;

Alignment Scores: Pred. No.: 1.82e+03 Length: 493  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 5 Gaps: 0

US-10-014-101B-41 (1-6) x ABV07966 (1-493)

QY 1 IleThrArgAlaArglle 6  
 DB 98 ATTACCCGTGCCGGATA 115

RESULT 8

AAC36788  
 ID AAC36788 standard; DNA; 503 BP.

XX AC AAC36788;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 15080.

KW Hybridisation assay; genetic mapping; gene expression control;  
 KW protein identification; signal transduction pathway; metabolic pathway;  
 KW promoter; termination sequence; ss.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-00301439.

XX PR 25-FEB-1999; 99US-0121825P.

XX PR 05-MAR-1999; 99US-0123180P.

XX PR 09-MAR-1999; 99US-0123548P.

XX PR 23-MAR-1999; 99US-0125788P.

XX PR 25-MAR-1999; 99US-0126264P.

XX PR 29-MAR-1999; 99US-0126785P.

XX PR 01-APR-1999; 99US-0127462P.

XX PR 08-APR-1999; 99US-0128234P.

XX PR 06-APR-1999; 99US-0128714P.

XX PR 16-APR-1999; 99US-0129845P.

XX PR 19-APR-1999; 99US-0130077P.

XX PR 21-APR-1999; 99US-0130449P.

XX PR 23-APR-1999; 99US-0130510P.

XX PR 28-APR-1999; 99US-0131449P.

XX PR 30-APR-1999; 99US-0132048P.

XX PR 04-MAY-1999; 99US-0132407P.

XX PR 05-MAY-1999; 99US-0132484P.

XX PR 06-MAY-1999; 99US-0132485P.

XX PR 06-MAY-1999; 99US-0132486P.

XX PR 07-MAY-1999; 99US-0132487P.

XX PR 11-MAY-1999; 99US-0132863P.

XX PR 14-MAY-1999; 99US-0134256P.

XX PR 14-MAY-1999; 99US-0134218P.

XX PR 14-MAY-1999; 99US-0134219P.

XX PR 14-MAY-1999; 99US-0134221P.

us-10-014-101b-41.rng

Tue Feb 22 10:19:08 2005

PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 21-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 23-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152263P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159337P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160747P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.

PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161922P.  
 PR 28-OCT-1999; 99US-0161992P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:  
 Pred. No.: 1.86e+03 Length: 503  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 3 Gaps: 0

US-10-014-101B-41 (1-6) x AAC36788 (1-503)

QY 1 IleThrArgAlaArgile 6  
 DB 104 ATAACGAGACCGAGAATT 121

RESULT 9  
 ABQ24115  
 ID ABQ24115 standard; DNA; 537 BP.

XX AC ABQ24115;

XX DT 12-JUL-2002 (first entry)

XX DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 10706.

XX KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;  
 KW drug; side effect; cancer; central nervous system; cardiovascular;  
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;  
 KW SNP; cell differentiation; ds.

XX OS Homo sapiens.

XX PN WO200218632-A2.

XX PD 07-MAR-2002.

XX PF 01-SEP-2001; 2001WO-EP010074.

XX PR 01-SEP-2000; 2000DE-01043826.

XX PR 05-SEP-2000; 2000DE-01044543.

XX PA (EPIG-) EPIGENOMICS AG.

XX PI Olek A, Piepenbrock C, Berlin K, Guetig D;

XX DR WPI; 2002-371829/40.

XX PT Determining the degree of cytosine methylation in genomic DNA, useful for  
 PT diagnosis and prognosis, comprises selective hybridization of amplicons  
 PT from chemically treated DNA.

XX PS Claim 12; 56pp + Sequence Listing; 56pp; German.

XX CC This invention describes a novel method for determining the degree of  
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
 CC genomic sample of DNA. The sample is treated chemically to convert  
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic  
 CC DNA that contains the target C is amplified to form a labeled amplicon.  
 CC The amplicon is hybridised to two classes, each with at least one member,  
 CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the  
 CC degree of hybridisation to both classes is determined from the label on  
 CC the amplicon. From the ratio of labels hybridised to the two classes of  
 CC oligomers, the degree of methylation is calculated. The method is used:  
 CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs  
 CC and of a wide range of diseases, e.g. cancer, disorders of the central

CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,  
 CC particularly by detecting mutations or single nucleotide polymorphisms  
 CC (SNP/s); and (ii) for differentiation of cell or tissue types and for  
 CC investigating cell differentiation. The method allows the methylation  
 CC status of many C residues to be determined simultaneously. ABQ13410-  
 CC ABQ54121 represent genomic DNA sequences used to illustrate the method  
 CC for determining the degree of cytosine methylation described in the  
 CC disclosure of the invention

XX SQ Sequence 537 BP; 192 A; 204 C; 49 G; 92 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 2e+03 Length: 537  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-10-014-101B-41 (1-6) x ABQ24115 (1-537)

QY 1 IleThrArgAlaArgile 6  
 DB 240 ATAACCCGCGCAGGAATA 257

RESULT 10  
 ABQ24114/c  
 ID ABQ24114 standard; DNA; 537 BP.

XX AC ABQ24114;

XX DT 12-JUL-2002 (first entry)

XX DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 10705.

XX KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;  
 KW drug; side effect; cancer; central nervous system; cardiovascular;  
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;  
 KW SNP; cell differentiation; ds.

XX OS Homo sapiens.

XX PN WO200218632-A2.

XX PD 07-MAR-2002.

XX PF 01-SEP-2001; 2001WO-EP010074.

XX PR 01-SEP-2000; 2000DE-01043826.

XX PR 05-SEP-2000; 2000DE-01044543.

XX PA (EPIG-) EPIGENOMICS AG.

XX PI Olek A, Piepenbrock C, Berlin K, Guetig D;

XX DR WPI; 2002-371829/40.

XX PT Determining the degree of cytosine methylation in genomic DNA, useful for  
 PT diagnosis and prognosis, comprises selective hybridization of amplicons  
 PT from chemically treated DNA.

XX PS Claim 12; 56pp + Sequence Listing; 56pp; German.

XX CC This invention describes a novel method for determining the degree of  
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
 CC genomic sample of DNA. The sample is treated chemically to convert  
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic  
 CC DNA that contains the target C is amplified to form a labeled amplicon.  
 CC The amplicon is hybridised to two classes, each with at least one member,  
 CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the  
 CC degree of hybridisation to both classes is determined from the label on  
 CC the amplicon. From the ratio of labels hybridised to the two classes of  
 CC oligomers, the degree of methylation is calculated. The method is used:  
 CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs  
 CC and of a wide range of diseases, e.g. cancer, disorders of the central

CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs  
 CC and of a wide range of diseases, e.g. cancer, disorders of the central  
 CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,  
 CC particularly by detecting mutations or single nucleotide polymorphisms  
 CC (SNP's); and (ii) for differentiation of cell or tissue types and for  
 CC investigating cell differentiation. The method allows the methylation  
 CC status of many C residues to be determined simultaneously. ABQ13410-  
 CC ABQ54121 represent genomic DNA sequences used to illustrate the method  
 CC for determining the degree of cytosine methylation described in the  
 CC disclosure of the invention  
 XX

SQ Sequence 537 BP; 92 A; 49 C; 204 G; 192 T; 0 U; 0 Other;

Alignment Scores: Length: 537  
 Pred. No.: 2e+03 Matches: 6  
 Score: 27.00  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-10-014-101B-41 (1-6) x ABQ24114 (1-537)

QY 1 IleThrArgAlaArgile 6  
 |||||  
 DB 298 ATACCCGCGCAGGATA 281

RESULT 11

ACA21667  
 ID ACA21667 standard; DNA; 708 BP.

AC ACA21667;

DT 19-JUN-2003 (first entry)

DE Prokaryotic essential gene #3324.

Antisense; ds; prokaryotic essential gene; cell proliferation;  
 drug design; gene.

Bacillus anthracis.

WO200277183-A2.

03-OCT-2002.

21-MAR-2002; 2002WO-US009107.

21-MAR-2001; 2001US-00815242.

06-SEP-2001; 2001US-00948993.

25-OCT-2001; 2001US-0342923P.

08-FEB-2002; 2002US-00072851.

06-MAR-2002; 2002US-0362699P.

(ELIT-) ELITRA PHARM INC.

Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

WPI; 2003-029926/02.

P-ESDB; ABU17797.

New antisense nucleic acids, useful for identifying proteins or screening  
 for homologous nucleic acids required for cellular proliferation to  
 isolate candidate molecules for rational drug discovery programs.

Claim 14; SEQ ID NO 9537; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of  
 the 6213 antisense sequences given in the specification where expression  
 of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
 CC prokaryotic essential genes. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

SQ Sequence 708 BP; 273 A; 85 C; 155 G; 195 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2.71e+03 Length: 708  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 8 Gaps: 0

US-10-014-101B-41 (1-6) x ACA21667 (1-708)

QY 1 IleThrArgAlaArgile 6  
 |||||  
 DB 150 ATTACAGGCGGAGGATA 167

RESULT 12

AAQ99805

ID AAQ99805 standard; cDNA; 831 BP.

AC AAQ99805;

DT 20-JUN-1996 (first entry)

DE Thaumatin like gene PR-5mz.

SAR; tobacco; protein-synthesis independent gene; cyclohexamide;

systemic acquired resistance response; anti-pathogen; plant protection;  
 maize; PR-5; ss.

Zea mays.

Key Location/Qualifiers

CDS 72..590

FT /\*tag= a

XX MO9519443-A2.

PN 20-JUL-1995.

XX 03-JAN-1995; 95WO-IB000002.

XX 13-JAN-1994; 94US-00181271.

PA (CIBA ) CIBA GEIGY AG.

PI Ryals JA, Alexander DC, Uknes SJ, Ward ER;  
 XX WPI; 1995-263872/34.  
 DR P-PSDB; AAR81082.  
 XX  
 PT New DNA contg. plant systemic acquired resistance genes - and transgenic  
 PT plants contg. them, impart disease and pest resistance, also Arabidopsis  
 PT gene promoter to control DNA transcription.  
 XX  
 PS Claim 21; Page 72-73; 85pp; English.  
 XX  
 CC This sequence represents the DNA sequence of a maize thaumatin like gene,  
 CC PR-5mz. This sequence was isolated by screening a BTH-induced cDNA  
 CC library of maize, with a probe matching a known rice thaumatin clone.  
 CC This sequence, and AAQ9800-Q99804 are all used in recombinant/chimaeric  
 CC DNA molecules of the invention. These sequences were isolated by  
 CC differential screening of a cDNA library, followed by analysis by  
 CC Northern hybridisation to RNA in the presence and absence of  
 CC cyclohexamide. The genes are used in the creation of transgenic plants.  
 CC All of these sequences confer anti-pathogenic properties to transgenic  
 CC plants. Transgenic expression of 2 or more of the recombinant molecules  
 CC of the invention that encode anti-pathogenic proteins provides a  
 CC synergistic increase in plant protection, and may also offer protection  
 CC against a wider range of pathogens  
 XX  
 SQ Sequence 831 BP; 182 A; 259 C; 230 G; 160 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 3.23e+03 Length: 831  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0  
 US-10-014-101B-41 (1-6) x AAQ99805 (1-831)  
 QY 1 IleThrArgAlaArgIle 6  
 DB 684 ATTACGTCACGCAPT 701  
 RESULT 13  
 AAC69410  
 ID AAC69410 standard; cDNA; 874 BP.  
 XX  
 AC AAC69410;  
 AC  
 DT 30-JAN-2001 (first entry)  
 XX  
 DE Human secreted protein gene 12 SEQ ID NO:22.  
 XX  
 KW Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;  
 KW antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;  
 KW cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide;  
 KW fungicide; ophthalmological; gene therapy; autoimmune disease; infection;  
 KW hyperproliferative disorder; cardiovascular disorder; angiogenesis;  
 KW cerebrovascular disorder; nervous system disorder; ocular disorder;  
 KW wound healing; skin aging; food additive; preservative; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200058468-A2.  
 XX  
 PD 05-OCT-2000.  
 XX  
 PF 22-MAR-2000; 2000WO-US007526.  
 XX  
 PR 26-MAR-1999; 99US-0126600P.  
 PR 22-DEC-1999; 99US-0171550P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Ruben SM, Komatsoulis G;

XX WPI; 2000-611713/58.  
 DR P-PSDB; AAB38130.  
 XX  
 PT Nucleic acids encoding human secreted proteins, used to prevent, treat,  
 PT ameliorate, or diagnose conditions such as autoimmune disorders, skin  
 PT disorders and cancer.  
 XX  
 PS Claim 1; Page 323-324; 374pp; English.  
 XX  
 CC The polynucleotide sequences given in AAC69399 to AAC69445 encode the  
 CC human secreted proteins given in AAB38119 to AAB38165. AAB38166 to  
 CC AAB38201 represent human secreted polypeptide sequences and proteins  
 CC homologous to them, which are given in the exemplification of the present  
 CC invention. Human secreted proteins have activities based on the tissues  
 CC and cells the genes are expressed in. Example of activities include:  
 CC immunosuppressive; antiarthritic; antirheumatic; antiproliferative;  
 CC cytostatic; cardiant; vasotropic; cerebroprotective; neurotropic;  
 CC neuroprotective; antibacterial; virucide; fungicide; and  
 CC ophthalmological. The polynucleotides and polypeptides can be used to  
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used  
 CC in diagnosing a pathological condition or susceptibility to a  
 CC pathological condition. Disorders which are diagnosed or treated include  
 CC autoimmune diseases, hyperproliferative disorders, cardiovascular  
 CC disorders, cerebrovascular disorders, angiogenesis, nervous system  
 CC disorders, infections caused by bacteria, viruses and fungi and ocular  
 CC disorders. The polypeptides can also be used to aid wound healing and  
 CC epithelial cell proliferation, to prevent skin aging due to sunburn, to  
 CC maintain organs before transplantation, for supporting cell culture of  
 CC primary tissues, to regenerate tissues and in chemotaxis. The  
 CC polypeptides can also be used as a food additive or preservative to  
 CC increase or decrease storage capabilities. AAC69399 to AAC69398 and  
 CC AAB38118 represent sequences used in the exemplification of the present  
 CC invention  
 XX  
 SQ Sequence 874 BP; 237 A; 216 C; 232 G; 185 T; 0 U; 4 Other;  
 Alignment Scores:  
 Pred. No.: 3.42e+03 Length: 874  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 3 Gaps: 0  
 US-10-014-101B-41 (1-6) x AAC69410 (1-874)  
 QY 1 IleThrArgAlaArgIle 6  
 DB 717 ATAACAAGGGCTAGAATA 734  
 RESULT 14  
 AAV32472  
 ID AAV32472 standard; DNA; 900 BP.  
 XX  
 AC AAV32472;  
 AC  
 DT 17-OCT-2003 (revised)  
 DT 10-SEP-1998 (first entry)  
 XX  
 DE Ubiquitin gene of Mamestra brassicae nuclear polyhedrosis virus.  
 XX  
 KW Ubiquitin gene; ds.  
 XX  
 OS Mamestra brassicae nucleopolyhedrovirus.  
 XX  
 FH Key Location/Qualifiers  
 FT 5'UTR 1..405 /\*tag= b  
 FT promoter 1..40 /\*tag= a  
 FT CDS 406..708

P-PSDB; ABO79930.

Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide, useful as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from bacterial infection.

Disclosure; SEQ ID NO 12105; 455pp; English.

The invention relates to Pseudomonas aeruginosa polypeptides and the polynucleotides encoding them. The sequences are useful in diagnosis and therapy of pathological conditions, as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from a bacterial infection, for evaluating a compound, such as a polypeptide, for the ability to bind a P. aeruginosa nucleic acid, as components of effective antibacterial targets, as targets for antibacterial drugs, including anti-P. aeruginosa drugs, as templates for recombinant production of P. aeruginosa-derived peptides or polypeptides, as target components for diagnosis and/or treatment of P. aeruginosa-caused infection, and in detection of P. aeruginosa sequences or other sequences of Pseudomonas species using biochip technology. Sequences ABD01397-ABD17967 represent P. aeruginosa polynucleotides of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html

Sequence 936 BP; 151 A; 321 C; 319 G; 145 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3-68e+03	Length:	936
Score:	27,00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	11	Gaps:	0

US-10-014-101B-41 (1-6) x ABD13501 (1-936)

QY 1 IleThrArgAlaArgIle 6  
|||||  
Db 333 ATCACCAGGCGGAGGATC 316

RESULT 16  
ADG32343  
ID ADG32343 standard; DNA; 984 BP.  
XX AC ADG32343;  
XX AC  
XX DT 26-FEB-2004 (first entry)  
XX DE DNA encoding the precursor fusion protein AFP RS-scfV CWPD2 SeqID 52.  
XX KW radish; scfV; ds; anti-fungal peptide; AFP; scfV; disease resistant;  
KW transgenic; plant; fungal infection; antibody; pathogen-specific antibody; fungicidal; agricultural; chimeric; chicken; gene.  
XX OS Chimeric.  
OS Synthetic.  
OS Raphanus sativus.  
OS Gallus gallus.  
XX WO2003089475-A2.  
XX FN  
XX PD 30-OCT-2003.  
XX PF 14-APR-2003; 2003WO-EP003852.  
XX PR 22-APR-2002; 2002EP-00008929.  
PR 28-MAY-2002; 2002EP-00011807.  
XX (FEAU ) FRAUNHOFER GES FOERDERUNG ANGEWANDTEN.  
XX Peschen D, Fischer R, Schillberg S, Liao Y, Dorfmueller S;  
PI WPI; 2003-615309/58.

3'UTR /tag= C  
product= "Ubiquitin"  
709..900  
/tag= d

JP10150989-A.  
PN 09-JUN-1998.  
XX 26-NOV-1996; 96JP-00315069.  
XX 26-NOV-1996; 96JP-00315069.  
PA (NOYA-) NOVAKU BIOTECHNOLOGY KAIHATSU.  
XX WI; 1998-379992/33.  
DR P-PSDB; AA#48856.  
XX DNA coding ubiquitin from Mamestra brassicae nuclear polyhedrosis virus - used to detect Mamestra brassicae nuclear polyhedrosis virus.  
XX Claim 2; Page 5-6; 7pp; Japanese.  
XX The invention claims for the Mamestra brassicae nuclear polyhedrosis virus ubiquitin gene and the protein it encodes. The ubiquitin gene can be used as an index to detect the Mamestra brassicae nuclear polyhedrosis virus. (Updated on 17-OCT-2003 to standardise OS field)

Sequence 900 BP; 301 A; 175 C; 171 G; 253 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3-53e+03	Length:	900
Score:	27,00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-10-014-101B-41 (1-6) x AAV32472 (1-900)

QY 1 IleThrArgAlaArgIle 6  
|||||  
Db 656 ATAACCCGGCTCGAATA 673

RESULT 15  
ABD13501/c  
ID ABD13501 standard; DNA; 936 BP.  
XX AC ABD13501;  
XX DT 29-JUL-2004 (first entry)  
XX Pseudomonas aeruginosa polynucleotide #12105.  
DE Bacterial infection; gene; ds; Pseudomonas aeruginosa infection; antibacterial.  
KW Pseudomonas aeruginosa.  
XX OS US6551795-B1.  
XX FN  
XX PD 22-APR-2003.  
XX PF 18-FEB-1999; 99US-00252991.  
XX PR 18-FEB-1998; 98US-0074788P.  
PR 27-JUL-1998; 98US-0094190P.  
XX (GENO-) GENOME THERAPEUTICS CORP.  
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
FI WPI; 2003-615309/58.

XX WPI; 2003-854088/79.  
 DR P-PSDB; ADG32354.  
 XX  
 PT New fusion protein comprising an anti-fungal protein or peptide and an  
 PT antibody fragment, useful in agriculture and horticulture for producing  
 PT Ascomyceta-resistant transgenic plants, plant cells or plant tissues.  
 XX  
 PS Example 4; SEQ ID NO 52; 47pp; English.  
 XX  
 CC This invention relates to a novel fusion protein comprising an anti-  
 CC fungal protein or peptide (AFP) and an antibody fragment (scFv).  
 CC Specifically, it refers to APs recognizing an epitope of an Ascomyceta  
 CC and a cellular targeting sequence, which can be used to generate disease  
 CC resistant transgenic plants that are protected against fungal infection.  
 CC Accordingly, a method is described for antibody based resistance in  
 CC plants such that the undesirable and expensive chemical controls often  
 CC used in agriculture are not required. The present invention provides  
 CC antibodies, recombinant antibodies and fragments thereof, as well as  
 CC fusion proteins that can be used as pathogen-specific antibodies targeted  
 CC to different plant cell compartments. As such, these fungicidal agents  
 CC confer a broad spectrum of disease resistance in both economically  
 CC important crops and ornamental plants. This polynucleotide is a DNA  
 CC sequence encoding a precursor fusion protein of the order [AFP - linker -  
 CC antibody fragment] of the invention.  
 XX  
 SQ Sequence 984 BP; 204 A; 270 C; 310 G; 200 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 3.89e+03 Length: 984  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 10 Gaps: 0

US-10-014-101B-41 (1-6) x ADG32343 (1-984)

Qy 1 IleThrArgAlaArgIle 6  
 |||||  
 Db 59 ATACGGGTGCAGAAATC 76

RESULT 17

ADG32350  
 ID ADG32350 standard; DNA; 987 BP.

XX AC ADG32350;

XX 26-FEB-2004 (first entry)

XX DNA encoding the precursor fusion protein of AFP RS-scfV VDM2 SeqID 59.

XX radish; scFv; ds; anti-fungal peptide; AFP; scFv; disease resistant;  
 XX transgenic; plant; fungal infection; antibody;  
 KW pathogen-specific antibody; fungicidal; agriculture; mouse; chimeric;  
 KW murine; gene.

XX Chimeric.

OS Synthetic.

OS Raphanus sativus.

OS Mus musculus.

XX WO2003089475-A2.

XX 30-OCT-2003.

XX 14-APR-2003; 2003WO-EP003852.

XX 22-APR-2002; 2002EP-00008929.

XX 28-MAY-2002; 2002EP-00011807.

XX (FRAU ) FRAUNHOFER GES FOERDERUNG ANGEWANDTEN.

XX

PI Peschen D, Fischer R, Schillberg S, Liao Y, Dorfmueller S;  
 XX WPI; 2003-854088/79.  
 DR P-PSDB; ADG32361.  
 XX  
 XX New fusion protein comprising an anti-fungal protein or peptide and an  
 PT antibody fragment, useful in agriculture and horticulture for producing  
 PT Ascomyceta-resistant transgenic plants, plant cells or plant tissues.  
 XX  
 PS Example 15; SEQ ID NO 59; 47pp; English.  
 XX  
 CC This invention relates to a novel fusion protein comprising an anti-  
 CC fungal protein or peptide (AFP) and an antibody fragment (scFv).  
 CC Specifically, it refers to APs recognizing an epitope of an Ascomyceta  
 CC and a cellular targeting sequence, which can be used to generate disease  
 CC resistant transgenic plants that are protected against fungal infection.  
 CC Accordingly, a method is described for antibody based resistance in  
 CC plants such that the undesirable and expensive chemical controls often  
 CC used in agriculture are not required. The present invention provides  
 CC antibodies, recombinant antibodies and fragments thereof, as well as  
 CC fusion proteins that can be used as pathogen-specific antibodies targeted  
 CC to different plant cell compartments. As such, these fungicidal agents  
 CC confer a broad spectrum of disease resistance in both economically  
 CC important crops and ornamental plants. This polynucleotide is a DNA  
 CC sequence encoding a precursor fusion protein of the order [AFP - linker -  
 CC antibody fragment] of the invention.  
 XX

SQ Sequence 987 BP; 252 A; 225 C; 268 G; 242 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 3.9e+03 Length: 987  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 10 Gaps: 0

US-10-014-101B-41 (1-6) x ADG32350 (1-987)

Qy 1 IleThrArgAlaArgIle 6  
 |||||  
 Db 59 ATACGGGTGCAGAAATC 76

RESULT 18

ADG32348

ID ADG32348 standard; DNA; 987 BP.

XX AC ADG32348;

XX 26-FEB-2004 (first entry)

XX DNA encoding the precursor fusion protein of AFP RS-scfV VDCw SeqID 57.

XX radish; scFv; ds; anti-fungal peptide; AFP; scFv; disease resistant;  
 KW transgenic; plant; fungal infection; antibody;  
 KW pathogen-specific antibody; fungicidal; agriculture; mouse; chimeric;  
 KW murine; gene.

OS Chimeric.

OS Synthetic.

OS Raphanus sativus.

OS Mus musculus.

XX WO2003089475-A2.

XX 30-OCT-2003.

XX 14-APR-2003; 2003WO-EP003852.

XX 22-APR-2002; 2002EP-00008929.

XX 28-MAY-2002; 2002EP-00011807.

XX (FRAU ) FRAUNHOFER GES FOERDERUNG ANGEWANDTEN.

XX Peschen D, Fischer R, Schillberg S, Liao Y, Dorfmueller S;  
PI WPI; 2003-854088/79.  
XX P-PSDB; ADG32359.  
DR  
XX New fusion protein comprising an anti-fungal protein or peptide and an  
PT antibody fragment, useful in agriculture and horticulture for producing  
PT Ascomyceta-resistant transgenic plants, plant cells or plant tissues.  
XX  
PS Example 15; SEQ ID NO 57; 47pp; English.  
XX  
CC This invention relates to a novel fusion protein comprising an anti-  
CC fungal protein or peptide (AFP) and an antibody fragment (scFv).  
CC Specifically, it refers to AFPs recognizing an epitope of an Ascomyceta  
CC and a cellular targeting sequence, which can be used to generate disease  
CC resistant transgenic plants that are protected against fungal infection.  
CC Accordingly, a method is described for antibody based resistance in  
CC plants such that the undesirable and expensive chemical controls often  
CC used in agriculture are not required. The present invention provides  
CC antibodies, recombinant antibodies and fragments thereof, as well as  
CC fusion proteins that can be used as pathogen-specific antibodies targeted  
CC to different plant cell compartments. As such, these fungicidal agents  
CC confer a broad spectrum of disease resistance in both economically  
CC important crops and ornamental plants. This polynucleotide is a DNA  
CC sequence encoding a precursor fusion protein of the order [AFP - linker -  
CC antibody fragment] of the invention.  
XX  
SQ Sequence 987 BP; 249 A; 228 C; 271 G; 239 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 3.9e+03 Length: 987  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0  
US-10-014-101B-41 (1-6) x ADG32348 (1-987)  
QY 1 IleThrArgAlaArgIle 6  
DB 59 ATACCGGTGCAGATC 76  
RESULT 19  
ABV29237  
ID ABV29237 standard; cdna; 1039 BP.  
XX  
AC ABV29237;  
XX  
XX 16-SEP-2002 (first entry)  
XX Human prostate expression marker cdna 29228.  
DE  
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
XX pharmacogenomic marker; gene; ss.  
XX Homo sapiens.  
XX WO200160860-A2.  
XX  
PD 23-AUG-2001.  
XX  
PF 20-FEB-2001; 2001WO-US005171.  
XX  
PR 17-FEB-2000; 2000US-0183119P.  
PR 16-MAR-2000; 2000US-0189862P.  
PR 25-MAY-2000; 2000US-0207454P.  
PR 09-JUN-2000; 2000US-0211314P.  
PR 18-JUL-2000; 2000US-0219007P.  
PR 13-DEC-2000; 2000US-0255281P.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;  
XX WPI; 2001-662795/76.  
XX  
PT Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer.  
XX  
PS Claim 1; Page 6234-6235; 11750pp; English.  
XX  
CC The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for: (a) assessing whether  
CC a patient is afflicted with prostate cancer; (b) monitoring the  
CC progression of prostate cancer in a patient; (c) assessing the efficacy  
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or indolence of prostate cancer in a patient  
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX  
SQ Sequence 1039 BP; 294 A; 276 C; 280 G; 188 T; 0 U; 1 Other;  
Alignment Scores:  
Pred. No.: 4.13e+03 Length: 1039  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 5 Gaps: 0  
US-10-014-101B-41 (1-6) x ABV29237 (1-1039)  
QY 1 IleThrArgAlaArgIle 6  
DB 892 ATTACGGGCGCAGAATA 909  
RESULT 20  
ADC24030  
ID ADC24030 standard; DNA; 1059 BP.  
XX  
AC ADC24030;  
XX  
XX 18-DEC-2003 (first entry)  
DT  
XX DNA sequence (SeqID 297) encoding a nitrilase enzyme.  
DE  
XX gene; ds; nitrilase; nitrile; cyanohydrin; ammonia; biocatalyst;  
KW enantiomer; chiral medicine.  
XX  
OS Unidentified.  
XX WO2003000840-A2.  
PN  
PD 03-JAN-2003.  
XX  
XX 15-MAY-2002; 2002WO-US015983.  
PF  
XX 21-JUN-2001; 2001US-0300189P.  
PR 30-JUL-2001; 2001US-030906P.  
PR 22-JAN-2002; 2002US-0351336P.  
XX  
PA (DIVE-) DIVERSA CORP.  
PA (MADD/) MADDEN D.  
XX  
XX Madden M, Desantis G, Chaplin JA, Weiner DP, Milan A, Chi E;  
PI Short JW, Burk M;  
XX WPI; 2003-201417/19.  
DR P-PSDB; ADC24031.

XX Novel nitrilase polypeptide, useful for making (R)- or (S)-ethyl-4-cyano-  
PT 3-hydroxybutyric acid or (R)- or (S)-mandelic acid or (S)- or (R)-phenyl  
PT lactic acid derivative and for producing pharmaceutical composition, and  
PT food additive.  
XX  
PS Claim 1; SEQ ID NO 297; 560pp; English.  
XX  
CC This invention relates to nitrilases and the nucleic acids that encode  
CC these enzymes thereof. Specifically, it refers to polypeptides that  
CC exhibit nitrilase activity, i.e. the ability to directly hydrolyse  
CC nitriles or cyanohydrins into their corresponding carboxylic acids and  
CC ammonia. Nitrilases have commercial utility as biocatalysts for use in  
CC the synthesis of enantiomerically pure aromatic and aliphatic amino  
CC acids, as well as hydroxy acids, which are important for the development  
CC of chiral medicines. Furthermore, the present invention describes  
CC nitrilases, isolated from mesophilic microorganisms, that have improved  
CC activity and stability at increased pH and temperature. They are also  
CC inexpensive, efficient catalysts, have broad substrate specificity and  
CC are capable of chiral differentiation. This polynucleotide is a DNA  
CC sequence that encodes a nitrilase enzyme of the invention.  
XX  
SQ Sequence 1059 BP; 202 A; 322 C; 346 G; 189 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 4.22e+03 Length: 1059  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0  
US-10-014-101B-41 (1-6) x ADC24030 (1-1059)  
QY 1 IleThrArgAlaArgile 6  
Db 884 ATTACTCGCGCCGGATA 901  
RESULT 21  
ADH36131  
ID ADH36131 standard; DNA; 1059 BP.  
XX  
AC ADH36131;  
XX  
DT 11-MAR-2004 (first entry)  
XX  
DE Chemical process monitoring-related nitrilase gene sequence SeqID297.  
XX  
DE Chemical process monitoring; biochemical process monitoring; cyanide;  
KW high throughput system; gene; ds.  
XX  
OS Unidentified.  
XX  
XX WO2003098187-A2.  
XX  
XX 27-NOV-2003.  
XX  
XX PF 15-MAY-2003; 2003WO-US015639.  
XX  
XX PR 15-MAY-2002; 2002US-0380737P.  
XX  
XX PA (DIVE-) DIVERSA CORP.  
XX  
XX Weiner D, Chaplin JA, Chi E, Milan A, Desantis G, Burk MJ;  
PI McQuaid J, Stege J;  
XX  
XX WPI; 2004-142708/14.  
DR P-PSDB; ADH36132.  
XX  
XX Monitoring a chemical or biochemical process comprises providing a  
PT reactant comprising a cyanide or a material that can be converted to  
PT cyanide or a reactant that generates a cyanide or a material that can be  
PT converted to cyanide.

XX Claim 74; SEQ ID NO 297; 277pp; English.  
XX  
CC This invention relates to a novel method of monitoring chemical or  
CC biochemical processes. The method involves providing a reactant  
CC comprising cyanide (or a material that can be converted to a cyanide)  
CC that generates as a reaction product cyanide or a material that can be  
CC converted to cyanide and measuring the concentration of produced cyanide.  
CC The method is useful for monitoring a chemical or biochemical process.  
CC The method is effective for high throughput systems and is sufficiently  
CC sensitive to detect a small amount of product. The present sequence is  
CC that of a gene which encodes a nitrilase enzyme which can be used in the  
CC method of the invention.  
XX  
SQ Sequence 1059 BP; 202 A; 322 C; 346 G; 189 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 4.22e+03 Length: 1059  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0  
US-10-014-101B-41 (1-6) x ADH36131 (1-1059)  
QY 1 IleThrArgAlaArgile 6  
Db 884 ATTACTCGCGCCGGATA 901  
RESULT 22  
ADG93832  
ID ADG93832 standard; DNA; 1059 BP.  
XX  
AC ADG93832;  
XX  
DT 11-MAR-2004 (first entry)  
XX  
DE Nitrilase enzyme gene sequence SeqID297.  
XX  
KW nitrilase; nitrile; carboxylic acid; chemical process; pH; temperature;  
KW enantioselective transformation; gene; ds.  
XX  
OS Unidentified.  
XX  
XX WO2003097810-A2.  
XX  
XX 27-NOV-2003.  
XX  
XX PF 15-MAY-2003; 2003WO-US015712.  
XX  
XX PR 15-MAY-2002; 2002US-00146772.  
XX  
XX PR 09-SEP-2002; 2002US-00241742.  
XX  
XX PA (DIVE-) DIVERSA CORP.  
XX  
XX Desantis G, Short JM, Burk M, Wong K, Farwell R, Chatman K;  
PI WPI; 2004-090637/09.  
XX  
XX P-PSDB; ADG93833.  
XX  
XX New isolated or recombinant nucleic acid encoding a polypeptide having  
PT nitrilase activity, useful for screening enantioselective transformation.  
XX  
XX Claim 44; SEQ ID NO 297; 295pp; English.  
XX  
CC This invention is related to a novel isolated or recombinant nucleic acid  
CC encoding a protein having nitrilase activity. Nitrilase's are capable of  
CC converting nitrile's directly to carboxylic acids and have great  
CC potential for use in industrial chemical processes. The isolated  
CC nitrilase proteins of the invention have increased activity and stability  
CC at increased pH and temperature when compared to those conventionally  
CC used. In addition, the nucleic acid of the invention is useful for

CC screening enantioselective transformation. The present sequence is that  
 CC of a DNA sequence which encodes a nitrilase enzyme of the invention.

XX  
 SQ Sequence 1059 BP; 202 A; 322 C; 346 G; 189 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 4.22e+03 Length: 1059  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 12 Gaps: 0

US-10-014-101B-41 (1-6) x ADG93832 (1-1059)

OY 1 IleThrArgAlaArgIle 6  
 |||||  
 Db 884 ATTACTCGCGCCGGATA 901

RESULT 23

ADI62429  
 ID ADI62429 standard; DNA; 1059 BP.

XX  
 AC ADI62429;

DT 22-APR-2004 (first entry)

DE DNA encoding nitrilase polypeptide #149.

XX Atorvastatin; (R)-ethyl 4-cyano-3-hydroxybutyrate;  
 KW (R)-ethyl 4-cyano-3-hydroxybutyric acid; epichlorohydrin;  
 KW 3-hydroxyglutaronitrile; 3-hydroxyglutaronitrile;  
 KW 4-cyano-3-hydroxybutyric acid; ethyl-4-cyano-3-hydroxybutyric acid;  
 KW mixed hyperlipidaemia; homozygous familial hypercholesterolaemia;  
 KW antilipaeic; gene; ds.

XX Unidentified.

XX WO2003106415-A2.

PN 24-DEC-2003.

XX 13-JUN-2003; 2003WO-US018840.

XX 13-JUN-2002; 2002US-0389317P.

PR 28-JUN-2002; 2002US-0392944P.

XX (DIVE-) DIVERSA CORP.

PA Burk M, Desantis G, Morgan B, Zhu Z;

XX WPI; 2004-090821/09.

DR P-PSDB; ADI62430.

XX Preparation of atorvastatin comprises catalytic conversion of 3-  
 PT hydroxyglutaronitrile by polypeptide with nitrilase activity, converting  
 PT obtained 4-cyano-3-hydroxybutyric acid to ethyl-4-cyano-3-hydroxybutyric  
 PT acid and forming atorvastatin.

XX Claim 46; SEQ ID NO 297; 253pp; English.

XX The present invention relates to a method for preparing an atorvastatin  
 CC intermediate known as (R)-ethyl 4-cyano-3-hydroxybutyrate ((R)-ethyl 4-  
 CC cyano-3-hydroxybutyric acid). The method comprises optionally converting  
 CC epichlorohydrin or equivalent to 3-hydroxyglutaronitrile, catalytic  
 CC conversion of 3-hydroxyglutaronitrile or equivalent to 4-cyano-3-  
 CC hydroxybutyric acid with a polypeptide having nitrilase activity,  
 CC converting 4-cyano-3-hydroxybutyric acid to ethyl-4-cyano-3-hydroxybutyric  
 CC acid, and converting this to (R)-ethyl 4-cyano-3-hydroxybutyrate. The  
 CC method involves whole cell processes, cell lysate process, "one pot"  
 CC processes, and "multi-pot" processes using a variety of parameters.  
 CC Atorvastatin is used, in conjunction with dietary restriction, in the  
 CC management of hyperlipidaemia, including hypercholesterolaemia, mixed

CC dyslipidaemia and homozygous familial hypercholesterolaemia. The present  
 CC sequence encodes a nitrilase polypeptide obtained from an environmental  
 CC sample.

XX Sequence 1059 BP; 202 A; 322 C; 346 G; 189 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 4.22e+03 Length: 1059  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 12 Gaps: 0

US-10-014-101B-41 (1-6) x ADI62429 (1-1059)

OY 1 IleThrArgAlaArgIle 6  
 |||||  
 Db 884 ATTACTCGCGCCGGATA 901

RESULT 24

ADI64550  
 ID ADI64550 standard; DNA; 1059 BP.

XX  
 AC ADI64550;

DT 22-APR-2004 (first entry)

DE DNA encoding nitrilase seq id 149.

XX (R)-ethyl 4-cyano-3-hydroxybutyric acid; nitrile hydrolysis;  
 KW carboxylic acid; cyanohydrin moiety hydrolysis;  
 KW aminonitrile moiety hydrolysis; chiral alpha-hydroxy acid molecule;  
 KW chiral amino acid molecule; (S)-ethyl 4-cyano-3-hydroxybutyric acid;  
 KW (R)-mandelic acid; (S)-mandelic acid; (S)-phenyl lactic acid derivative;  
 KW (R)-phenyl lactic acid derivative; & enantiomeric excess;  
 KW & diastomeric excess; food additive; drug intermediate; ds; nitrilase;  
 KW gene.

XX Unidentified.

XX US2004014195-A1.

XX 22-JAN-2004.

XX 15-MAY-2003; 2003US-00440523.

XX 29-DEC-1999; 99US-0173609P.

PR 07-DEC-2000; 2000US-0254414P.

PR 28-DEC-2000; 2000US-00751299.

PR 21-JUN-2001; 2001US-0300189P.

PR 30-JUL-2001; 2001US-0309006P.

PR 22-JAN-2002; 2002US-0351336P.

PR 15-MAY-2002; 2002US-00146772.

PR 09-SEP-2002; 2002US-00241742.

XX (DIVE-) DIVERSA CORP.

XX Desantis G, Short JM, Burk MJ, Wong K, Farwell R, Chatman K;

XX WPI; 2004-121569/12.

DR P-PSDB; ADI64551.

XX Novel isolated or recombinant polypeptide having nitrilase activity,  
 PT useful in production of food additives.

XX Claim 1; SEQ ID NO 297; 105pp; English.

XX The invention describes an isolated or recombinant polypeptide (I)  
 CC comprising amino acids having a sequence at least 50 % identical to a  
 CC sequence (SI) available in electronic form (EC) from the following web  
 CC site ftp.segdata.uspto.gov/sequence.html?docID=2004014195, or its  
 CC variants, having one or more mutations at residue 55 Lys, Gly or Glu, at

CC residue 60 glutamic acid, at residue 111 Ser, their combinations or  
 CC fragments. (I) is useful for: producing an (R)-ethyl 4-cyano-3-  
 CC hydroxybutyric acid; hydrolysing a nitrile to a carboxylic acid;  
 CC hydrolysing cyanohydrin moiety or an aminonitrile moiety; producing a  
 CC chiral alpha-hydroxy acid molecule or a chiral amino acid molecule;  
 CC producing an (S)-ethyl 4-cyano-3-hydroxybutyric acid; producing an (R)-  
 CC mandelic acid or (S)-mandelic acid; producing (S)-phenyl lactic acid  
 CC derivative or an (R)-phenyl lactic acid derivative; modifying a molecule;  
 CC and for identifying a modified compound. The inventive method is useful  
 CC for monitoring or determining \* enantiomeric excess or % diastereomeric  
 CC excess. (I) is useful in the production of food additives and drug  
 CC intermediates. This sequence encodes a nitrilase of the invention.  
 XX  
 SQ Sequence 1059 BP; 202 A; 322 C; 346 G; 189 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 4.22e+03 Length: 1059  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 12 Gaps: 0

US-10-014-101B-41 (1-6) x ADI64550 (1-1059)

QY 1 IleThrArgAlaArglle 6  
 DB 884 ATTACTGCGCGCGGATA 901  
 RESULT 25  
 ADS60562/c  
 ID ADS60562 standard; cDNA; 1124 BP.  
 XX AC ADS60562;  
 XX 02-DEC-2004 (first entry)  
 DE Bacterial polynucleotide #12549.  
 XX Recombinant DNA construct; transformed plant; improved plant property;  
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
 KW pathogen tolerance; pest tolerance; plant disease resistance;  
 KW call cycle pathway modification; plant growth regulator;  
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
 KW bacterial polynucleotide; gene; ss.  
 XX Bacteria.  
 OS  
 XX US2003233675-A1.  
 XX 18-DEC-2003.  
 PD  
 XX 20-FEB-2003; 2003US-00369493.  
 PF  
 XX 21-FEB-2002; 2002US-0360039P.  
 PR  
 XX (CAOY/) CAO Y.  
 PA (HINK/) HINKLE G J.  
 PA (SLAT/) SLATER S C.  
 PA (CHEN/) CHEN X.  
 PA (GOLD/) GOLDMAN B S.  
 XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
 XX WPI; 2004-061375/06.  
 XX New recombinant DNA construct comprising a promoter positioned to provide  
 PT for expression of a polynucleotide encoding a polypeptide from a  
 PT microbial source, useful for producing plants with improved properties.  
 PS Claim 1; SEQ ID NO 36236; 122pp; English.  
 XX

CC The invention relates to a recombinant DNA construct comprising a  
 CC promoter functional in a plant cell, where the promoter is positioned to  
 CC provide for expression of a polynucleotide encoding a polypeptide from a  
 CC microbial source. The invention also relates to a transformed plant  
 CC comprising the recombinant DNA construct and a method of producing a  
 CC transformed plant having an improved property. The plant is a crop plant  
 CC such as maize or soybean. The method of producing a transformed plant  
 CC having an improved property comprises transforming a plant with the  
 CC recombinant DNA construct and growing the transformed plant, where the  
 CC polynucleotide or polypeptide is useful for improving plant properties.  
 CC The recombinant DNA construct is useful for producing plants with  
 CC improved plant properties, e.g. improved cold heat or drought tolerance,  
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
 CC increased resistance to plant disease, better growth rate by modification  
 CC of the cell cycle pathway with plant growth regulators, increased rate of  
 CC homologous recombination, modified seed oil or protein yield and/or  
 CC content, improved yield by modification of carbohydrate, nitrogen or  
 CC phosphorus use and/or uptake, by modification of photosynthesis or by  
 CC providing improved plant growth and development under at least one stress  
 CC condition, improved lignin production or improved galactomannan  
 CC production. This sequence represents a bacterial polynucleotide used in  
 CC the scope of the invention. Note: The sequence data for this patent did  
 CC not form part of the printed specification but was obtained in electronic  
 CC format from USPTO at seqdata.uspto.gov/sequence.html.  
 XX  
 SQ Sequence 1124 BP; 260 A; 326 C; 284 G; 254 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 4.5e+03 Length: 1124  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 13 Gaps: 0

US-10-014-101B-41 (1-6) x ADS60562 (1-1124)

QY 1 IleThrArgAlaArglle 6  
 DB 1010 ATCCTAGAGCGCGAATC 993  
 RESULT 26  
 ACA21136/c  
 ID ACA21136 standard; DNA; 1171 BP.  
 XX AC ACA21136;  
 XX 19-JUN-2003 (first entry)  
 DT  
 XX Prokaryotic essential gene #2793.  
 DE  
 XX Antisense; ds; prokaryotic essential gene; cell proliferation;  
 KW drug design; gene.  
 KW Acinetobacter baumannii.  
 OS  
 XX WO200277183-A2.  
 XX 03-OCT-2002.  
 PD  
 XX 21-MAR-2002; 2002WO-US009107.  
 PF  
 XX 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.  
 XX (ELIT-) ELITRA PHARM INC.  
 PA Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX

DR WPI: 2003-029926/02.  
DR P-PSDB; ABU17266.  
XX  
PT New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
PS Claim 14; SEQ ID NO 9006; 1766pp; English.  
XX  
CC The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
CC prokaryotic essential genes. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 1171 BP; 368 A; 146 C; 165 G; 492 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 4.71e+03 Length: 1171  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0

US-10-014-101B-41 (1-6) x ACA21136 (1-1171)  
QY 1 IleThrArgAlaArglie 6  
DB 422 ATTACAGAGCTCGGATC 405

RESULT 27  
ACA01962/c  
ID ACA01962 standard; DNA; 1215 BP.  
XX  
AC ACA01962;  
XX  
DT 04-JUN-2003 (first entry)  
XX  
DE C. glutamicum derived ORF SEQ ID 1953.  
XX  
KW Coryneform; nucleic acid array; fermentation; culture; ds.  
XX Corynebacterium glutamicum.  
OS  
PI DE10128510-A1.  
XX

PD 19-DEC-2002.  
XX  
PF 13-JUN-2001; 2001DE-01028510.  
XX  
PR 13-JUN-2001; 2001DE-01028510.  
XX  
PA (DEGS ) DEGUSSA AG.  
XX  
PI Farwick M, Moeckel B, Pfefferle W, Bathe B, Huthmacher K;  
XX  
DR WPI: 2003-279970/28.  
XX  
PT New nucleic acid array useful for monitoring mRNA expression of  
PT Corynebacterium glutamicum during fermentation, comprising nucleic acid  
PT from Corynebacterium glutamicum.  
XX  
PS Claim 1; Page 645-646; 709pp; German.  
XX  
CC This invention describes a novel nucleic acid array involving  
CC Corynebacterium glutamicum polynucleotides. The arrays are used to  
CC analyse C. glutamicum, particularly for monitoring a fermentation process  
CC to determine expression levels of C. glutamicum cellular mRNA. Such  
CC monitoring particularly differentiates between expression levels of  
CC different strains of C. glutamicum and allows the adjustment of different  
CC culture and fermentation conditions. ACA0010-ACA02188 represent C.  
CC glutamicum derived polynucleotides described in the disclosure of the  
CC invention  
XX  
SQ Sequence 1215 BP; 319 A; 248 C; 298 G; 350 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 4.91e+03 Length: 1215  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0

US-10-014-101B-41 (1-6) x ACA01962 (1-1215)  
QY 1 IleThrArgAlaArglie 6  
DB 284 ATCACAGGGCAGCATC 267

RESULT 28  
ABD13808  
ID ABD13808 standard; DNA; 1290 BP.  
XX  
AC ABD13808;  
XX  
DT 29-JUL-2004 (first entry)  
XX  
DE Pseudomonas aeruginosa polynucleotide #12412.  
XX  
KW Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;  
KW antibacterial.  
XX  
OS Pseudomonas aeruginosa.  
XX  
PN US6551795-B1.  
XX  
FD 22-APR-2003.  
XX  
PF 18-FEB-1999; 99US-00252991.  
XX  
PR 18-FEB-1998; 98US-0074788P.  
XX  
PR 27-JUL-1998; 98US-0094190P.  
XX  
PA (GENO-) GENOME THERAPEUTICS CORP.  
XX  
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
XX WPI: 2003-615309/58.  
XX

DR P-PSDB; ABO80237.  
 XX Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide,  
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
 PT pathological conditions resulting from bacterial infection.  
 XX Disclosure; SEQ ID NO 12412; 455pp; English.  
 XX The invention relates to *Pseudomonas aeruginosa* polypeptides and the  
 CC polynucleotides encoding them. The sequences are useful in diagnosis and  
 CC therapy of pathological conditions, as molecular targets for diagnostics,  
 CC prophylaxis and treatment of pathological conditions resulting from a  
 CC bacterial infection, for evaluating a compound, such as a polypeptide,  
 CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of  
 CC effective antibacterial targets, as targets for antibacterial drugs,  
 CC including anti-*P. aeruginosa* drugs, as templates for recombinant  
 CC production of *P. aeruginosa*-derived peptides or polypeptides, as target  
 CC components for diagnosis and/or treatment of *P. aeruginosa*-caused  
 CC infection, and in detection of *P. aeruginosa* sequences or other sequences  
 CC of *Pseudomonas* species using biochip technology. Sequences ABD01397-  
 CC ABD17967 represent *P. aeruginosa* polynucleotides of the invention. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification but was obtained in electronic format from USPTO at  
 CC seqdata.uspto.gov/sequence.html  
 XX SQ Sequence 1290 BP; 193 A; 434 C; 451 G; 212 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 5.24e+03 Length: 1290  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 11 Gaps: 0  
 US-10-014-101B-41 (1-6) x ABD13808 (1-1290)  
 QY 1 IleThrArgAlaArgile 6  
 DB 935 ATCACCAGGCGGAGGATC 952  
 RESULT 29  
 ABD13579/c  
 ID ABD13579 standard; DNA; 1311 BP.  
 XX AC ABD13579;  
 XX AC  
 XX 29-JUL-2004 (first entry)  
 XX *Pseudomonas aeruginosa* polynucleotide #12183.  
 DE Bacterial infection; gene; ds; *Pseudomonas aeruginosa* infection;  
 XX antibacterial.  
 XX *Pseudomonas aeruginosa*.  
 XX US6551795-B1.  
 XX 22-APR-2003.  
 XX 18-FEB-1999; 99US-00252991.  
 XX 18-FEB-1998; 98US-0074788P.  
 XX 27-JUL-1998; 98US-0094190P.  
 XX (GENO-) GENOME THERAPEUTICS CORP.  
 XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
 XX WPI; 2003-615309/58.  
 XX P-PSDB; ABO80008.  
 XX Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide,

PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
 PT pathological conditions resulting from bacterial infection.  
 XX Disclosure; SEQ ID NO 12183; 455pp; English.  
 XX The invention relates to *Pseudomonas aeruginosa* polypeptides and the  
 CC polynucleotides encoding them. The sequences are useful in diagnosis and  
 CC therapy of pathological conditions, as molecular targets for diagnostics,  
 CC prophylaxis and treatment of pathological conditions resulting from a  
 CC bacterial infection, for evaluating a compound, such as a polypeptide,  
 CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of  
 CC effective antibacterial targets, as targets for antibacterial drugs,  
 CC including anti-*P. aeruginosa* drugs, as templates for recombinant  
 CC production of *P. aeruginosa*-derived peptides or polypeptides, as target  
 CC components for diagnosis and/or treatment of *P. aeruginosa*-caused  
 CC infection, and in detection of *P. aeruginosa* sequences or other sequences  
 CC of *Pseudomonas* species using biochip technology. Sequences ABD01397-  
 CC ABD17967 represent *P. aeruginosa* polynucleotides of the invention. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification but was obtained in electronic format from USPTO at  
 CC seqdata.uspto.gov/sequence.html  
 XX SQ Sequence 1311 BP; 220 A; 456 C; 443 G; 192 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 5.33e+03 Length: 1311  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 11 Gaps: 0  
 US-10-014-101B-41 (1-6) x ABD13579 (1-1311)  
 QY 1 IleThrArgAlaArgile 6  
 DB 416 ATCACCAGGCGGAGGATC 399  
 RESULT 30  
 AAH65248/c  
 ID AAH65248 standard; DNA; 1434 BP.  
 XX AC AAH65248;  
 XX 26-SEP-2001 (first entry)  
 XX C glutamicum coding sequence fragment SEQ ID NO: 283.  
 XX *Corynebacterium*; amino acid synthesis; vitamin; saccharide;  
 XX organic acid synthesis; ds.  
 XX *Corynebacterium glutamicum*.  
 XX EF1108790-A2.  
 XX 20-JUN-2001.  
 XX 18-DEC-2000; 2000EP-00127688.  
 XX 16-DEC-1999; 99JP-00377484.  
 XX 07-APR-2000; 2000JP-00159162.  
 XX 03-AUG-2000; 2000JP-00280988.  
 XX (KYOW ) KYOWA HAKKO KOGYO KK.  
 XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;  
 XX Tateishi N, Senoh A, Ikeda M, Ozaki A;  
 XX WPI; 2001-376931/40.  
 XX P-PSDB; AAG90029.  
 XX Novel polynucleotides derived from *Corynebacterium*, for identifying  
 PT mutation point of a gene, measuring expression of a gene, analyzing

PT expression profile or pattern of a gene and identifying homologous gene.  
 XX Claim 8; SEQ ID NO 283; 246pp + Sequence Listing; English.  
 PS  
 XX The present invention provides a number of nucleotide and protein sequences from the *Corynebacterium glutamicum*. These are useful for identifying the mutation point of a gene derived from a mutant of *Corynebacterium glutamicum*, measuring expression amount and analysing the expression profile or expression pattern of a gene derived from *Corynebacterium glutamicum*, and identifying a homologue of a gene derived from *Corynebacterium glutamicum*. *Corynebacterium* bacteria are useful for producing amino acids, nucleic acids, vitamins, saccharides and organic acids, particularly L-lysine. The present sequence is a nucleic acid described in the exemplification of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the European Patent Office

XX  
 SQ Sequence 1434 BP; 381 A; 291 C; 360 G; 402 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 5.88e+03 Length: 1434  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 5 Indels: 0  
 DB: Gaps: 0

US-10-014-101B-41 (1-6) x AAH65248 (1-1434)

QY 1 IleThrArgAlaArgIle 6  
 |||||

Db 503 ATCAAGGGGACGCATC 486

RESULT 31  
 AAC44665  
 ID AAC44665 standard; DNA; 1436 BP.  
 XX  
 AC AAC44665;  
 XX  
 DT 18-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 43666.  
 XX  
 KW Hybridisation assay; Genetic mapping; gene expression control;  
 KW protein identification; signal transduction pathway; metabolic pathway;  
 KW promoter; termination sequence; ss.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN EPI033405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PF 25-FEB-2000; 2000EP-00301439.  
 XX  
 XX 25-FEB-1999; 99US-0121825P.  
 PR 05-MAR-1999; 99US-0123180P.  
 PR 09-MAR-1999; 99US-0123548P.  
 PR 23-MAR-1999; 99US-0125788P.  
 PR 25-MAR-1999; 99US-0126264P.  
 PR 29-MAR-1999; 99US-0126785P.  
 PR 01-APR-1999; 99US-0127462P.  
 PR 06-APR-1999; 99US-0128234P.  
 PR 08-APR-1999; 99US-0128714P.  
 PR 16-APR-1999; 99US-0129845P.  
 PR 19-APR-1999; 99US-0130077P.  
 PR 23-APR-1999; 99US-0130449P.  
 PR 23-APR-1999; 99US-0130510P.  
 PR 28-APR-1999; 99US-0130891P.  
 PR 30-APR-1999; 99US-0131449P.  
 PR 30-APR-1999; 99US-0132048P.  
 PR 04-MAY-1999; 99US-0132407P.  
 PR 99US-0132484P.

105-MAY-1999; 99US-0132485P.  
 06-MAY-1999; 99US-0132486P.  
 06-MAY-1999; 99US-0132487P.  
 07-MAY-1999; 99US-0132863P.  
 11-MAY-1999; 99US-0134256P.  
 14-MAY-1999; 99US-0134218P.  
 14-MAY-1999; 99US-0134219P.  
 14-MAY-1999; 99US-0134221P.  
 14-MAY-1999; 99US-0134370P.  
 18-MAY-1999; 99US-0134768P.  
 19-MAY-1999; 99US-0134941P.  
 20-MAY-1999; 99US-0135124P.  
 21-MAY-1999; 99US-0135353P.  
 24-MAY-1999; 99US-0135629P.  
 25-MAY-1999; 99US-0136021P.  
 27-MAY-1999; 99US-0136392P.  
 28-MAY-1999; 99US-0136782P.  
 01-JUN-1999; 99US-0137222P.  
 03-JUN-1999; 99US-0137528P.  
 04-JUN-1999; 99US-0137502P.  
 07-JUN-1999; 99US-0137724P.  
 08-JUN-1999; 99US-0138094P.  
 10-JUN-1999; 99US-0138540P.  
 10-JUN-1999; 99US-0138847P.  
 14-JUN-1999; 99US-0139119P.  
 16-JUN-1999; 99US-0139452P.  
 16-JUN-1999; 99US-0139453P.  
 17-JUN-1999; 99US-0139492P.  
 18-JUN-1999; 99US-0139454P.  
 18-JUN-1999; 99US-0139455P.  
 18-JUN-1999; 99US-0139456P.  
 18-JUN-1999; 99US-0139457P.  
 18-JUN-1999; 99US-0139458P.  
 18-JUN-1999; 99US-0139459P.  
 18-JUN-1999; 99US-0139460P.  
 18-JUN-1999; 99US-0139461P.  
 18-JUN-1999; 99US-0139462P.  
 18-JUN-1999; 99US-0139463P.  
 18-JUN-1999; 99US-0139750P.  
 18-JUN-1999; 99US-0139763P.  
 21-JUN-1999; 99US-0139817P.  
 22-JUN-1999; 99US-0139899P.  
 22-JUN-1999; 99US-0140353P.  
 23-JUN-1999; 99US-0140354P.  
 24-JUN-1999; 99US-0140695P.  
 28-JUN-1999; 99US-0140823P.  
 29-JUN-1999; 99US-0140991P.  
 30-JUN-1999; 99US-0141287P.  
 01-JUL-1999; 99US-0141842P.  
 01-JUL-1999; 99US-0142154P.  
 02-JUL-1999; 99US-0142055P.  
 06-JUL-1999; 99US-0142390P.  
 08-JUL-1999; 99US-0142803P.  
 09-JUL-1999; 99US-0142920P.  
 12-JUL-1999; 99US-0142977P.  
 13-JUL-1999; 99US-0143542P.  
 14-JUL-1999; 99US-0143624P.  
 15-JUL-1999; 99US-0144005P.  
 16-JUL-1999; 99US-0144085P.  
 16-JUL-1999; 99US-0144086P.  
 19-JUL-1999; 99US-0144325P.  
 19-JUL-1999; 99US-0144331P.  
 19-JUL-1999; 99US-0144332P.  
 19-JUL-1999; 99US-0144333P.  
 19-JUL-1999; 99US-0144334P.  
 19-JUL-1999; 99US-0144335P.  
 20-JUL-1999; 99US-0144352P.  
 20-JUL-1999; 99US-0144632P.  
 20-JUL-1999; 99US-0144884P.  
 21-JUL-1999; 99US-0144814P.  
 21-JUL-1999; 99US-0145086P.  
 21-JUL-1999; 99US-0145088P.  
 22-JUL-1999; 99US-0145085P.

```
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145092P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145313P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 08-AUG-1999; 99US-0147416P.
PR 08-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 28-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.

PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161320P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 5,89e+03
Score: 27.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 3

US-10-014-101b-41 (1-6) x AAC44665 (1-1436)

QY 1 IleThrArgAlaArgile 6
Db 1087 ATTACAGGGCGGAGGATT 1104

RESULT 32
ABK28625
ID ABK28625 standard; cdna; 1506 BP.
XX
AC ABK28625;
XX
DT 09-APR-2002 (first entry)
XX
DE cdna encoding A. thaliana cytokinin oxidase AtCKX2.
DE
KW Arabidopsis thaliana; plant; cytokinin oxidase; transgenic plant;
KW root growth; lateral root; adventitious root; root geotropism; herbicide;
KW root meristem; shoot meristem; leaf senescence; parthenocarp; gene; ss.
XX
OS Arabidopsis thaliana.
XX
XX WO200196580-A2.
XX
PD 20-DEC-2001.
XX
XX 18-JUN-2001; 2001WO-EP006833.
XX
PR 16-JUN-2000; 2000EP-00870132.
PR 27-DEC-2000; 2000US-0258415P.
PR 16-MAR-2001; 2001EP-00870053.
XX
XX (SCHW/) SCHMULLING T.
XX (WERN/) WERNER T.
XX
PI Schmullling T, Werner T;
XX
DR WPI; 2002-130736/17.
XX
PT Polynucleotide encoding novel plant protein having cytokinin oxidase
PT activity and the protein useful for stimulating root growth, enhancing
PT the formation of lateral or adventitious roots, altering root geotropism.
XX
PS Claim 2; Page 146-147; 154pp; English.
XX
CC The invention relates to an isolated polynucleotide (I) encoding a novel
CC plant protein (II) having cytokinin oxidase activity. (I) is useful for
CC production of transgenic plants, plant cells or tissues; for production
```

of altered plants, plant cell or tissues; and for effecting the expression of (II) where (I) is operably linked to one or more control sequences. The methods further comprises regenerating a plant from the plant cell. (I) and (II) are useful for stimulating root growth; enhancing the formation of lateral or adventitious roots; altering root geotropism, leading to an increase in yield; and for screening growth promoting chemical of herbicides. (I) is useful for increasing the size of the root meristem; increasing root size; increasing the size of the shoot meristem; delaying leaf senescence and altering leaf senescence; increasing leaf thickness; reducing or increasing the vessel size; inducing parthenocarpy; improving standability of the seedlings; increasing branching and for improving lodging resistance. Antibody (III) to (II) is useful for identifying and obtaining proteins interacting with (II) comprising a screening assay, preferably a two-hybrid screening assay. ABK28606-ABK28633 represent Arabidopsis thaliana cytokinin oxidase coding sequences and PCR primers of the invention

SQ Sequence 1506 BP; 431 A; 328 C; 324 G; 423 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	6.21e+03	Length:	1506
Score:	27.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-014-101B-41 (1-6) x ABK28625 (1-1506)

QY 1 IleThrArgAlaArgIle 6  
 |||||  
 646 ATAACGAGAGCCAGAAATT 663

RESULT 33

ACC85294  
 ID ACC85294 standard; cDNA; 1506 BP.

XX ACC85294;

DT 18-SEP-2003 (first entry)

DE Arabidopsis cytokinin oxidase-like protein 2 cDNA.

KW Maize; root growth; root geotropism; cytokinin oxidase; seed size;  
 KW embryo size; cotyledon size; transgenic plant; herbicide; gene; ss.

XX Arabidopsis thaliana.

OS WO2003050287-A2.

XX 19-JUN-2003.

PF 10-DEC-2002; 2002WO-EP013990.

XX 10-DEC-2001; 2001US-00014101.

XX (SCHM/) SCHMULLING T.

PA (WERN/) WERNER T.

XX Schmullling T, Werner T;

XX WPI; 2003-541577/51.

XX Stimulating root growth, enhancing lateral or adventitious root formation  
 PT or altering root geotropism comprises increasing plant cytokinin oxidase  
 PT levels or other protein or nucleic acid that reduces active cytokinins in  
 PT a plant.

PS Claim 3; Page 167-168; 177pp; English.

XX The present invention relates to a method for stimulating root growth or  
 CC enhancing the formation of lateral or adventitious roots or altering root  
 CC geotropism, which comprises increasing in a plant or plant part the level

of a plant cytokinin oxidase or other protein that reduces the level of active cytokinins in a plant or plant part. Cytokinin oxidase protein and coding sequences from Arabidopsis thaliana are also provided. The method is useful in modifying plant morphological, biochemical and physiological properties, such as in modifying the initiation, stimulation or enhancement of root growth, adventitious root formation, lateral root formation, root geotropism, shoot growth, apical dominance, branching, timing of senescence, timing of flowering, flower formation, seed development and/or seed yield. The present sequence is a coding sequence shown in the invention

SQ Sequence 1506 BP; 431 A; 328 C; 324 G; 423 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	6.21e+03	Length:	1506
Score:	27.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-10-014-101B-41 (1-6) x ACC85294 (1-1506)

QY 1 IleThrArgAlaArgIle 6  
 |||||  
 646 ATAACGAGAGCCAGAAATT 663

RESULT 34

ADH61267  
 ID ADH61267 standard; DNA; 1506 BP.

XX ADH61267;

DT 25-MAR-2004 (first entry)

XX Arabidopsis thaliana cytokinin oxidase (CKX1) DNA #1.

KW Male-sterility; cytokinin oxidase; transgenic plant; seed production;  
 KW CKX1; mouse-ear cress; gene; ds.

XX Arabidopsis thaliana.

XX US2003163847-A1.

XX 28-AUG-2003.

XX 20-DEC-2002; 2002US-00326184.

XX 20-DEC-2001; 2001US-0343129P.

XX (PHAA ) PHARMACIA CORP.

XX Huang S, Crossland LD, Cheikh N, Morris RO;

XX WPI; 2003-897983/82.

XX GENBANK; AF303978.

XX Producing plants characterized by reversible male-sterility, useful for  
 PT maintaining male sterility in plants, by transforming a plant cell with a  
 PT nucleic acid construct containing a polynucleotide encoding a cytokinin  
 PT oxidase.

XX Claim 41; SEQ ID NO 8; 33pp; English.

XX The invention relates to a method for producing a plant characterised by  
 CC reversible male-sterility which involves transforming a plant cell with a  
 CC nucleic acid construct containing a polynucleotide encoding a cytokinin  
 CC oxidase. The method is useful for producing reversible male-sterility in  
 CC transgenic plants, or for maintaining male sterility in plants. The  
 CC method reduces the expense of seed production for existing hybrid plants  
 CC such as corn, but also makes it possible to produce hybrid varieties of  
 CC traditionally non-hybrid crops. The method is also useful for introducing  
 CC economically valuable traits from plants having undesirable production

CC characteristics into plants having desirable characteristics. The present  
 CC sequence is Arabidopsis thaliana cytokinin oxidase (CKX1) DNA. This  
 CC sequence is used to illustrate the method of the invention.

XX  
 SQ Sequence 1506 BP; 431 A; 329 C; 323 G; 423 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 6.21e+03 Length: 1506  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 10 Gaps: 0

US-10-014-101B-41 (1-6) x ADH61267 (1-1506)

Qy 1 IleThrArgAlaArgile 6  
 |||||  
 Db 646 ATACGAGAGCCAGATT 663

RESULT 35

AD006498  
 ID AD006498 standard; DNA; 1506 BP.

AC AD006498;

29-JUL-2004 (first entry)

A thaliana cytokinin oxidase AtCKX2 coding sequence.

bioremediation; AtCKX2; cytokinin availability; contaminant;  
 metal deficiency; nutrition; ds; gene; cytokinin oxidase.

Arabidopsis thaliana.

Key Location/Qualifiers  
 CDS 1..1506  
 /\*tag= a  
 /product= "AtCKX2"

WO2004038027-A1.

06-MAY-2004.

24-OCT-2003; 2003WO-EP012051.

24-OCT-2002; 2002EP-00079481.

(CROP-) CROPEDESIGN NV.

Van Camp W;

WPI; 2004-375913/35.

P-PSDB; AD006499.

Method for bioremediation, useful for removing contaminants or metals, by  
 decreasing cytokinin availability in plants, and cultivating plant on  
 substrate comprising one or more contaminants.

Claim 8; Page 55; 61pp; English.

The present invention relates to a method for bioremediation, which  
 involves decreasing cytokinin availability in a plant relative to a  
 corresponding wild type plant, and cultivating the plant on a substrate  
 comprising one or more contaminants, or cultivating a plant having  
 lowered availability of cytokinin relative to corresponding wild type  
 plants, on a substrate, which is to be treated. The method is useful for  
 bioremediation, for concentration of contaminants in a plant, where the  
 plant has a higher concentration of contaminants compared to a  
 corresponding wild type plant. The metal contaminants include aluminum,  
 americium, antimony, arsenic, barium, beryllium, bismuth, cadmium,  
 caesium, cerium, chromium, copper, gallium, germanium, gold, indium,  
 iridium, iron, lead, manganese, mercury, molybdenum, neptunium, osmium,

CC palladium, platinum, plutonium, radium, rhenium, rhodium, rubidium,  
 CC ruthenium, scandium, selenium, silver, strontium, technetium, tellurium,  
 CC thallium, tin, tungsten, uranium, vanadium or yttrium, preferably  
 CC cadmium. The plants obtained by the method is useful in bioremediation.  
 CC Transgenic plants with a lower availability of cytokinin are useful in  
 CC the manufacture of a medicament for treatment of disorders arising from  
 CC metal deficiencies, and as a medicament for improving animal or human  
 CC nutrition. The present sequence is a coding sequence shown in the  
 CC exemplification of the invention.

SQ Sequence 1506 BP; 431 A; 328 C; 324 G; 423 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 6.21e+03 Length: 1506  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 12 Gaps: 0

US-10-014-101B-41 (1-6) x AD006498 (1-1506)

Qy 1 IleThrArgAlaArgile 6  
 |||||  
 Db 646 ATACGAGAGCCAGATT 663

RESULT 36

ABK28629

ID ABK28629 standard; cDNA; 1515 BP.

AC ABK28629;

09-APR-2002 (first entry)

cDNA encoding A. thaliana cytokinin oxidase AtCKX6.

Arabidopsis thaliana; plant; cytokinin oxidase; transgenic plant;  
 root growth; lateral root; adventitious root; root geotropism; herbicide;  
 root meristem; shoot meristem; leaf senescence; parthenocarp; gene; ss.

Arabidopsis thaliana.

WO200196580-A2.

20-DEC-2001.

18-JUN-2001; 2001WO-EP006833.

16-JUN-2000; 2000EP-00870132.

27-DEC-2000; 2000US-0258415P.

16-MAR-2001; 2001EP-00870053.

(SCHM/) SCHMULLING T.

(WERN/) WERNER T.

Schmulling T, Werner T;

WPI; 2002-130736/17.

Polynucleotide encoding novel plant protein having cytokinin oxidase  
 activity and the protein useful for stimulating root growth, enhancing  
 the formation of lateral or adventitious roots, altering root geotropism.

Example 1; Page 149; 154pp; English.

The invention relates to an isolated polynucleotide (I) encoding a novel  
 plant protein (II) having cytokinin oxidase activity. (I) is useful for  
 production of transgenic plants, plant cells or tissues; for production  
 of altered plants, plant cell or tissues; and for effecting the  
 expression of (II) where (I) is operably linked to one or more control  
 sequences. The methods further comprises regenerating a plant from the  
 plant cell. (I) and (II) are useful for stimulating root growth;  
 enhancing the formation of lateral or adventitious roots; altering root

CC geotropism, leading to an increase in yield; and for screening growth  
 CC promoting chemical of herbicides. (I) is useful for increasing the size  
 CC of the root meristem; increasing root size; increasing the size of the  
 CC shoot meristem; delaying leaf senescence and altering leaf senescence;  
 CC increasing leaf thickness; reducing or increasing the vessel size;  
 CC inducing parthenocarp; improving standability of the seedlings;  
 CC to (II) is useful for identifying and obtaining proteins interacting with  
 CC (III) comprising a screening assay, preferably a two-hybrid screening  
 CC assay. ABK28608-ABK28633 represent Arabidopsis thaliana cytokinin oxidase  
 CC coding sequences and PCR primers of the invention  
 XX  
 SQ Sequence 1515 BP; 456 A; 327 C; 378 G; 354 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 6.25e+03 Length: 1515  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-10-014-101B-41 (1-6) x ABK28629 (1-1515)

QY 1 IleThrArgAlaArgile 6  
 DB 673 ATACCGCGGCAAGAATA 690

# RESULT 37

ACC85298  
 ID ACC85298 standard; cDNA; 1515 BP.

AC ACC85298;

DT 18-SEP-2003 (first entry)

XX Arabidopsis cytokinin oxidase-like protein 6 cDNA.

XX Maize; root growth; root geotropism; cytokinin oxidase; seed size;  
 XX embryo size; cotyledon size; transgenic plant; herbicide; gene; ss.

XX Arabidopsis thaliana.

XX WO2003050287-A2.

XX 19-JUN-2003.

XX 10-DEC-2002; 2002WO-EP013990.

XX 10-DEC-2001; 2001US-00014101.

XX (SCHM/) SCHMULLING T.

XX (WERN/) WERNER T.

XX Schmulling T, Werner T;

XX WPI; 2003-541577/51.

XX Stimulating root growth, enhancing lateral or adventitious root formation  
 PT or altering root geotropism comprises increasing plant cytokinin oxidase  
 PT levels or other protein or nucleic acid that reduces active cytokinins in  
 PT a plant.

PS Claim 2; Page 171-172; 177pp; English.

XX The present invention relates to a method for stimulating root growth or  
 CC enhancing the formation of lateral or adventitious roots or altering root  
 CC geotropism, which comprises increasing in a plant or plant part the level  
 CC of a plant cytokinin oxidase or other protein that reduces the level of  
 CC active cytokinins in a plant or plant part. Cytokinin oxidase protein and  
 CC coding sequences from Arabidopsis thaliana are also provided. The method  
 CC is useful in modifying plant morphological, biochemical and physiological  
 CC properties, such as in modifying the initiation, stimulation or

CC enhancement of root growth, adventitious root formation, lateral root  
 CC formation, root geotropism, shoot growth, apical dominance, branching,  
 CC timing of senescence, timing of flowering, flower formation, seed  
 CC development and/or seed yield. The present sequence is a coding sequence  
 CC shown in the invention

SQ Sequence 1515 BP; 456 A; 327 C; 378 G; 354 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 6.25e+03 Length: 1515  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-10-014-101B-41 (1-6) x ACC85298 (1-1515)

QY 1 IleThrArgAlaArgile 6

DB 673 ATACCGCGGCAAGAATA 690

# RESULT 38

ACA26612  
 ID ACA26612 standard; DNA; 1539 BP.

XX ACA26612;

XX 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #8269.

XX Antisense; ds; prokaryotic essential gene; cell proliferation;

XX drug design; gene.

XX Burkholderia mallei.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

XX P-PSDB; ABU22742.

XX New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.

XX Claim 14; SEQ ID NO 14482; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for

CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
CC prokaryotic essential genes. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 1539 BP; 197 A; 514 C; 605 G; 223 T; 0 U; 0 Other;

Alignment Scores:                      6.36e+03                      Length:                      1539  
Pred. No.:                      27.00                      Matches:                      6  
Score:                      100.00%                      Conservative:                      0  
Percent Similarity:                      100.00%                      Mismatches:                      0  
Best Local Similarity:                      100.00%                      Indels:                      0  
Query Match:                      100.00%                      Gaps:                      0  
DB:                      8

US-10-014-101b-41 (1-6) x ACA26612 (1-1539)

OY                      1 IleThrArgAlaArgIle 6  
                    |||||                    |||||  
Db                      1136 ATCACACGCGCGCATC 1153

RESULT 39

AAC43214  
ID AAC43214 standard; DNA; 1548 BP.

XX AAC43214;

XX DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 38443.

XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-00301439.

XX PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 16-APR-1999; 99US-0128714P.

PR 19-APR-1999; 99US-0129845P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0132863P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0136229P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.

R	20-JUL-1999	99US-0144884P
R	21-JUL-1999	99US-0144814P
R	21-JUL-1999	99US-0145086P
R	21-JUL-1999	99US-0145088P
R	22-JUL-1999	99US-0145085P
R	22-JUL-1999	99US-0145087P
R	22-JUL-1999	99US-0145089P
R	22-JUL-1999	99US-0145192P
R	22-JUL-1999	99US-0145145P
R	23-JUL-1999	99US-0145218P
R	23-JUL-1999	99US-0145224P
R	23-JUL-1999	99US-0145276P
R	26-JUL-1999	99US-0145913P
R	27-JUL-1999	99US-0145918P
R	27-JUL-1999	99US-0145919P
R	27-JUL-1999	99US-0145951P
R	28-JUL-1999	99US-0146386P
R	02-AUG-1999	99US-0146388P
R	02-AUG-1999	99US-0146389P
R	03-AUG-1999	99US-0147038P
R	04-AUG-1999	99US-0147204P
R	04-AUG-1999	99US-0147302P
R	05-AUG-1999	99US-0147192P
R	05-AUG-1999	99US-0147260P
R	06-AUG-1999	99US-0147303P
R	06-AUG-1999	99US-0147416P
R	09-AUG-1999	99US-0147493P
R	09-AUG-1999	99US-0147935P
R	10-AUG-1999	99US-0148171P
R	11-AUG-1999	99US-0148319P
R	12-AUG-1999	99US-0148341P
R	13-AUG-1999	99US-0148565P
R	13-AUG-1999	99US-0148684P
R	16-AUG-1999	99US-0149368P
R	17-AUG-1999	99US-0149175P
R	18-AUG-1999	99US-0149426P
R	20-AUG-1999	99US-0149722P
R	20-AUG-1999	99US-0149723P
R	20-AUG-1999	99US-0149923P
R	23-AUG-1999	99US-0149902P
R	23-AUG-1999	99US-0149930P
R	25-AUG-1999	99US-0150566P
R	26-AUG-1999	99US-0150884P
R	27-AUG-1999	99US-0151065P
R	27-AUG-1999	99US-0151068P
R	27-AUG-1999	99US-0151080P
R	30-AUG-1999	99US-0151303P
R	31-AUG-1999	99US-0151438P
R	01-SEP-1999	99US-0151930P
R	07-SEP-1999	99US-0152363P
R	10-SEP-1999	99US-0153070P
R	13-SEP-1999	99US-0153758P
R	15-SEP-1999	99US-0154018P
R	16-SEP-1999	99US-0154039P
R	20-SEP-1999	99US-0154779P
R	22-SEP-1999	99US-0155139P
R	23-SEP-1999	99US-0155486P
R	28-SEP-1999	99US-0155659P
R	24-SEP-1999	99US-0156458P
R	28-SEP-1999	99US-0158596P
R	04-OCT-1999	99US-0157117P
R	04-OCT-1999	99US-0157753P
R	06-OCT-1999	99US-0157865P
R	07-OCT-1999	99US-0158029P
R	08-OCT-1999	99US-0158232P
R	12-OCT-1999	99US-0158369P
R	13-OCT-1999	99US-0158929P
R	13-OCT-1999	99US-0159294P
R	13-OCT-1999	99US-0159325P
R	14-OCT-1999	99US-0159329P
R	14-OCT-1999	99US-0159330P
R	14-OCT-1999	99US-0159337P
R	14-OCT-1999	99US-0159342P

XX

CC This invention describes novel polynucleotides that encode protein  
 CC markers and fine chemical-production proteins from *Corynebacterium*  
 CC glutamicum. The polynucleotides are isolated from a nucleic acid library  
 CC of *C. glutamicum* then mutated at the specified positions, cloned and  
 CC expressed by standard methods. Cells, especially *Corynebacterium*  
 CC glutamicum, containing vectors that express the polynucleotides are used  
 CC for production of fine chemicals, preferably amino acids and specifically  
 CC lysine, but more generally nucleotides, nucleosides, lipids, fatty acids,  
 CC diols, carbohydrates, aromatic compounds, vitamins, co-factors and  
 CC enzymes. These are useful in the food, animal feed, cosmetics and  
 CC pharmaceutical industries. The polynucleotides, optionally as primers and  
 CC probes, can also be used for identification and classification of *C.*  
 CC glutamicum and related species, e.g. for diagnosis, for genomic mapping,  
 CC functional or evolutionary studies, gene manipulation and modulation of  
 CC metabolic activity. Cells that containing the polynucleotides of the  
 CC invention may produce fine chemicals in better yields, with higher  
 CC productivity and/or more efficiently. NOTE: This sequence is not  
 CC represented in the printed specification but is available in electronic  
 CC format. The sequence represented in this record has been obtained from  
 CC WO2003046123.

XX

SQ Sequence 1564 BP; 416 A; 314 C; 396 G; 438 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	6.47e+03	Length:	1564
Score:	27.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	11	Gaps:	0

US-10-014-101B-41 (1-6) x ADL66030 (1-1564)

QY 1 IleThrArgAlaArgIle 6

|||||

603 ATCACAGGGCGCAGCATC 586

DB

Search completed: February 18, 2005, 03:27:13  
 Job time : 112.16 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 18, 2005, 02:33:38 ; Search time 35.4 Seconds  
(without alignments)  
277.335 Million cell updates/sec

Title: US-10-014-101B-41

Perfect score: 27

Sequence: 1 ITRARI 6

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10014101/runat.16022005.075846.7970/app.query.fasta.1.796  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi  
-LIST=45 -DOCLALIGN=200 -THR SCORE=spect -THR MAX=100 -THR MIN=0 -ALIGN=40  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10014101 @CGN 1.1.124 @runat.16022005.075846.7970 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:\*

1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	27	100.0	936	4	US-09-252-991A-12105
C 2	27	100.0	1290	4	US-09-252-991A-12105
C 3	27	100.0	1311	4	US-09-252-991A-12183
C 4	27	100.0	1605	3	US-09-124-541-3
C 5	27	100.0	1605	4	US-09-663-326-3
C 6	27	100.0	1839	2	US-08-915-207-3
C 7	27	100.0	1839	3	US-09-238-555-3
C 8	27	100.0	1842	2	US-08-915-207-1
C 9	27	100.0	1842	3	US-09-238-555-1
C 10	27	100.0	2013	4	US-09-614-221A-614
C 11	27	100.0	2875	4	US-08-328-314-1
C 12	27	100.0	2875	1	US-08-731-045-1

13	27	100.0	6733	3	US-09-124-541-2	Sequence 2, Appli
14	27	100.0	6733	4	US-09-663-326-2	Sequence 2, Appli
C 15	27	100.0	11014	4	US-08-956-171E-91	Sequence 91, Appli
C 16	27	100.0	11014	4	US-08-781-986A-91	Sequence 91, Appli
C 17	27	100.0	29771	4	US-09-949-016-12754	Sequence 12754, A
18	27	100.0	29771	4	US-09-949-016-13956	Sequence 13956, A
19	27	100.0	70014	4	US-09-949-016-17110	Sequence 17110, A
C 20	27	100.0	141115	4	US-09-949-016-17490	Sequence 17490, A
C 21	26	96.3	75	3	US-08-671-824-20	Sequence 20, Appli
C 22	26	96.3	75	4	US-09-686-341-16	Sequence 16, Appli
C 23	26	96.3	185	4	US-09-313-294A-2688	Sequence 2688, Ap
C 24	26	96.3	283	4	US-09-313-294A-3534	Sequence 3534, Ap
C 25	26	96.3	291	4	US-09-313-294A-7329	Sequence 7329, Ap
C 26	26	96.3	297	4	US-09-313-294A-4423	Sequence 4423, Ap
C 27	26	96.3	473	4	US-09-270-767-5082	Sequence 5082, Ap
C 28	26	96.3	473	4	US-09-270-767-20364	Sequence 20364, A
C 29	26	96.3	601	4	US-09-949-016-33777	Sequence 33777, A
C 30	26	96.3	601	4	US-09-949-016-44357	Sequence 44357, A
C 31	26	96.3	601	4	US-09-949-016-44358	Sequence 44358, A
C 32	26	96.3	601	4	US-09-949-016-52245	Sequence 52245, A
C 33	26	96.3	601	4	US-09-949-016-52246	Sequence 52246, A
C 34	26	96.3	601	4	US-09-949-016-60727	Sequence 60727, A
C 35	26	96.3	601	4	US-09-949-016-110816	Sequence 110816,
C 36	26	96.3	601	4	US-09-949-016-110977	Sequence 110977,
C 37	26	96.3	601	4	US-09-949-016-111138	Sequence 111138,
C 38	26	96.3	601	4	US-09-949-016-111299	Sequence 111299,
C 39	26	96.3	714	3	US-08-998-416-867	Sequence 150894,
C 40	26	96.3	798	4	US-09-252-991A-7733	Sequence 867, App
C 41	26	96.3	1165	4	US-09-949-016-5024	Sequence 7733, Ap
C 42	26	96.3	1176	4	US-09-252-991A-11843	Sequence 5024, Ap
C 43	26	96.3	2073	3	US-09-221-017B-894	Sequence 11843, A
C 44	26	96.3	2103	4	US-09-252-991A-11773	Sequence 894, App
C 45	26	96.3	2103	4	US-09-252-991A-11773	Sequence 11773, A

ALIGNMENTS

RESULT 1

US-09-252-991A-12105/c  
; Sequence 12105, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 12105  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-12105

Alignment Scores:  
Pred. No.: 1.26e+03 Length: 936  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-41 (1-6) x US-09-252-991A-12105 (1-936)

QY 1 IleThrArgAlaArgile 6  
DB 333 ATCCAGGGCGAGGATC 316

## RESULT 2

US-09-252-991A-12412  
; Sequence 12412, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 12412  
; LENGTH: 1290  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-12412

Alignment Scores:  
Pred. No.: 1.76e+03 Length: 1290  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-41 (1-6) x US-09-252-991A-12412 (1-1290)

QY 1 IleThrArgAlaArgIle 6  
|||||  
Db 935 ATCACCAGGCGGAGGATC 952

## RESULT 3

US-09-252-991A-12183/c  
; Sequence 12183, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 12183  
; LENGTH: 1311  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-12183

Alignment Scores:  
Pred. No.: 1.79e+03 Length: 1311  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-41 (1-6) x US-09-252-991A-12183 (1-1311)

QY 1 IleThrArgAlaArgIle 6  
|||||  
Db 416 ATCACCAGGCGGAGGATC 399

## RESULT 4

US-09-124-541-3  
; Sequence 3, Application US/09124541A  
; Patent No. 6229066  
; GENERAL INFORMATION:  
; APPLICANT: Morris Ph.D., Roy O.  
; TITLE OF INVENTION: A CYTOKININ OXIDASE  
; FILE REFERENCE: UMO1490  
; CURRENT APPLICATION NUMBER: US/09/124,541A  
; CURRENT FILING DATE: 1998-07-29  
; EARLIER APPLICATION NUMBER: 60/054,268  
; EARLIER FILING DATE: 1997-07-30  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1605  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1605)  
US-09-124-541-3

Alignment Scores:  
Pred. No.: 2.21e+03 Length: 1605  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-10-014-101B-41 (1-6) x US-09-124-541-3 (1-1605)

QY 1 IleThrArgAlaArgIle 6  
|||||  
Db 703 ATCACCAGGCGGCGGATC 720

## RESULT 5

US-09-663-326-3  
; Sequence 3, Application US/09663326  
; Patent No. 6617497  
; GENERAL INFORMATION:  
; APPLICANT: Morris Ph.D., Roy O.  
; TITLE OF INVENTION: A CYTOKININ OXIDASE  
; FILE REFERENCE: UMO1490  
; CURRENT APPLICATION NUMBER: US/09/663,326  
; CURRENT FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: 60/054,268  
; PRIOR FILING DATE: 1997-07-30  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1605  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1605)  
US-09-663-326-3

Alignment Scores:  
Pred. No.: 2.21e+03 Length: 1605  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-41 (1-6) x US-09-663-326-3 (1-1605)

QY 1 IleThrArgAlaArgIle 6  
|||||  
Db 703 ATCACCAGGCGGCGGATC 720

```

RESULT 6
US-08-915-207-3
; Sequence 3, Application US/08915207
; Patent No. 5882897
; GENERAL INFORMATION:
; APPLICANT: Burnham, Martin K.R.
; APPLICANT: Lonetto, Michael A.
; APPLICANT: Warren, Patrick V.
; TITLE OF INVENTION: NOVEL YitJ
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,207
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Todd Q
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: P50549-3
; TELEPHONE: 215-994-2252
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1839 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-915-207-3
Alignment Scores:
Pred. No.: 2.54e+03 Length: 1839
Score: 27.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0
US-10-014-101B-41 (1-6) x US-08-915-207-3 (1-1839)
QY 1 IleThrArgAlaArgIle 6
Db 651 ATTACCAGACGACGCATT 668
RESULT 7
US-09-238-555-3
; Sequence 3, Application US/09238555
; Patent No. 6251652
; GENERAL INFORMATION:
; APPLICANT: Burnham, Martin K.R.
; APPLICANT: Lonetto, Michael A.
; APPLICANT: Warren, Patrick V.
; TITLE OF INVENTION: NOVEL YitJ
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia

```

```

; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/238,555
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/915,207
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Todd Q
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: P50549-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2252
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1839 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-09-238-555-3
Alignment Scores:
Pred. No.: 2.54e+03 Length: 1839
Score: 27.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0
US-10-014-101B-41 (1-6) x US-09-238-555-3 (1-1839)
QY 1 IleThrArgAlaArgIle 6
Db 651 ATTACCAGACGACGCATT 668
RESULT 8
US-08-915-207-1
; Sequence 1, Application US/08915207
; Patent No. 5882897
; GENERAL INFORMATION:
; APPLICANT: Burnham, Martin K.R.
; APPLICANT: Lonetto, Michael A.
; APPLICANT: Warren, Patrick V.
; TITLE OF INVENTION: NOVEL YitJ
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,207
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:

```

```
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Todd Q
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: P50549-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2252
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1842 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
; US-08-915-207-1
;
; Alignment Scores:
; Pred. No.: 2,55e+03 1842
; Score: 27.00 6
; Percent Similarity: 100.00% 0
; Best Local Similarity: 100.00% 0
; Query Match: 100.00% 0
; DB: 2
;
; US-10-014-101B-41 (1-6) x US-08-915-207-1 (1-1842)
;
; QY 1 lleThrArgAlaArgile 6
;
; Db 651 ATTACCAGACGCGCATT 668
;
; RESULT 9
; US-09-238-555-1
; Sequence 1, Application US/09238555
; Patent No. 6251652
; GENERAL INFORMATION:
; APPLICANT: Burnham, Martin K.R.
; APPLICANT: Lonetto, Michael A.
; APPLICANT: Warren, Patrick V.
; TITLE OF INVENTION: NOVEL YITJ
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/238,555
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/915,207
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Todd Q
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: P50549-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2252
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1842 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
```

```
; TOPOLOGY: linear
; US-09-238-555-1
;
; Alignment Scores:
; Pred. No.: 2,55e+03 1842
; Score: 27.00 6
; Percent Similarity: 100.00% 0
; Best Local Similarity: 100.00% 0
; Query Match: 100.00% 0
; DB: 3
;
; US-10-014-101B-41 (1-6) x US-09-238-555-1 (1-1842)
;
; QY 1 lleThrArgAlaArgile 6
;
; Db 651 ATTACCAGACGCGCATT 668
;
; RESULT 10
; US-09-614-221A-614
; Sequence 614, Application US/09614221A
; Patent No. 6723837
; GENERAL INFORMATION:
; APPLICANT: Karunanandaa, Balasulojini
; APPLICANT: Yu, Jaehyuk
; APPLICANT: Kishore, Ganesh M.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED
; FILE REFERENCE: 16516.075
; CURRENT APPLICATION NUMBER: US/09/614,221A
; CURRENT FILING DATE: 2000-07-12
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 626
; SEQ ID NO 614
; LENGTH: 2013
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
; US-09-614-221A-614
;
; Alignment Scores:
; Pred. No.: 2,79e+03 2013
; Score: 27.00 6
; Percent Similarity: 100.00% 0
; Best Local Similarity: 100.00% 0
; Query Match: 100.00% 0
; DB: 4
;
; US-10-014-101B-41 (1-6) x US-09-614-221A-614 (1-2013)
;
; QY 1 lleThrArgAlaArgile 6
;
; Db 228 ATTACCGTGGCCGAATT 245
;
; RESULT 11
; US-08-328-314-1
; Sequence 1, Application US/08328314
; Patent No. 5674728
; GENERAL INFORMATION:
; APPLICANT: Buxton, Frank
; APPLICANT: Jarai, Gabor
; APPLICANT: Visser, Jacob
; TITLE OF INVENTION: Fungal Protease
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII editor
; CURRENT APPLICATION NUMBER: US/08/328,314
; FILING DATE: TBA
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 4-19746/A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8615
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2875 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: pepe
; STRAIN: Aspergillus niger N400
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(1269..1370, 1462..1612, 1669..2323, 2382
; LOCATION: ..2667)
; OTHER INFORMATION: /function= "Aspartic Protease"
; OTHER INFORMATION: /product= "pepe"
; OTHER INFORMATION: /gene= "pepe"
; FEATURE:
; NAME/KEY: intron
; LOCATION: order(1371..1461, 1613..1668, 2324..2381)
; FEATURE:
; NAME/KEY: exon
; LOCATION: join(1269..1370, 1462..1612, 1669..2323, 2382
; LOCATION: ..2667)
; US-08-328-314-1
;
Alignment Scores:
Pred. No.: 4.05e+03 Length: 2875
Score: 27.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-10-014-101B-41 (1-6) x US-08-328-314-1 (1-2875)
QY 1 IleThrArgAlaArgile 6
Db 788 ATCACC CGCGCTCGGATT 805

RESULT 12
US-08-731-045-1
; Sequence 1, Application US/08731045
; Patent No. 5756338
; GENERAL INFORMATION:
; APPLICANT: Buxton, Frank
; APPLICANT: Jarai, Gabor
; APPLICANT: Visser, Jacob
; TITLE OF INVENTION: Fungal Protease
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY CORPORATION
; STREET: 520 White Plains Road
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591-9725
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Editor
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/731,045
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/328,314
; FILING DATE: October 24, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferraro, Gregory D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 4-19746/A/DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 277-3318
; TELEFAX: (908) 277-4306
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2875 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: pepe
; STRAIN: Aspergillus niger N400
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(1269..1370, 1462..1612, 1669..2323, 2382
; LOCATION: ..2667)
; OTHER INFORMATION: /function= "Aspartic Protease"
; OTHER INFORMATION: /product= "pepe"
; OTHER INFORMATION: /gene= "pepe"
; FEATURE:
; NAME/KEY: intron
; LOCATION: order(1371..1461, 1613..1668, 2324..2381)
; FEATURE:
; NAME/KEY: exon
; LOCATION: join(1269..1370, 1462..1612, 1669..2323, 2382
; LOCATION: ..2667)
; US-08-731-045-1
;
Alignment Scores:
Pred. No.: 4.05e+03 Length: 2875
Score: 27.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-10-014-101B-41 (1-6) x US-08-731-045-1 (1-2875)
QY 1 IleThrArgAlaArgile 6
Db 788 ATCACC CGCGCTCGGATT 805

RESULT 13
US-09-124-541-2
; Sequence 2, Application US/09124541A
; Patent No. 6229066
; GENERAL INFORMATION:
; APPLICANT: Morris Ph.D., Roy O.
; TITLE OF INVENTION: A CYTOKININ OXIDASE
; FILE REFERENCE: UMO1490
; CURRENT APPLICATION NUMBER: US/09/124,541A
; CURRENT FILING DATE: 1998-07-29
; EARLIER APPLICATION NUMBER: 60/054,268
; EARLIER FILING DATE: 1997-07-30
; NUMBER OF SEQ ID NOS: 20

```

```

; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 6733
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)..(6733)
; OTHER INFORMATION: genomic sequence for a cytokinin oxidase from Zea
; OTHER INFORMATION: mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1497)..(2111)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2524)..(3216)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3311)..(3607)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (5697)
; US-09-124-541-2

Alignment Scores:
Pred. No.: 9.8e+03 Length: 6733
Score: 27.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-014-101B-41 (1-6) x US-09-124-541-2 (1-6733)

Qy 1 IleThrArgAlaArgile 6
Db 2611 ATCACC GGCGCCGGATC 2628

RESULT 14
US-09-663-326-2
; Sequence 2, Application US/09663326
; Patent No. 6617497
; GENERAL INFORMATION:
; APPLICANT: Morris Ph.D., Roy O.
; TITLE OF INVENTION: A CYTOKININ OXIDASE
; FILE REFERENCE: UMO1490
; CURRENT APPLICATION NUMBER: US/09/663,326
; CURRENT FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 60/054,268
; PRIOR FILING DATE: 1997-07-30
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 6733
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)..(6733)
; OTHER INFORMATION: genomic sequence for a cytokinin oxidase from Zea
; OTHER INFORMATION: mays
; NAME/KEY: CDS
; LOCATION: (1497)..(2111)
; NAME/KEY: CDS
; LOCATION: (2524)..(3216)
; NAME/KEY: CDS
; LOCATION: (3311)..(3607)
; NAME/KEY: unsure
; LOCATION: (5697)
; US-09-663-326-2

Alignment Scores:
Pred. No.: 9.8e+03 Length: 6733
Score: 27.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-014-101B-41 (1-6) x US-09-663-326-2 (1-6733)

Qy 1 IleThrArgAlaArgile 6
Db 2611 ATCACC GGCGCCGGATC 2628

RESULT 15
US-08-956-171E-91/c
; Sequence 91, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11014 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 91:
US-08-956-171E-91

Alignment Scores:
Pred. No.: 1.63e+04 Length: 11014
Score: 27.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-014-101B-41 (1-6) x US-08-956-171E-91 (1-11014)

Qy 1 IleThrArgAlaArgile 6
```

Db 6555 ATTACGAGCAGCATT 6538

## RESULT 16

US-08-781-986A-91/c  
; Sequence 91, Application US/08781986A  
; Patent No. 6737248  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 5255  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/781,986A  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Benson, Bob  
; REGISTRATION NUMBER: 30,446  
; REFERENCE/DOCKET NUMBER: PB248PP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 91:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11014 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-781-986A-91

Alignment Scores: 11014  
Pred. No.: 1.63e+04  
Score: 27.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
Indels: 0  
Gaps: 0

US-10-014-101B-41 (1-6) x US-08-781-986A-91 (1-11014)

QY 1 IleThrArgAlaArgile 6

Db 6555 ATTACGAGCAGCATT 6538

## RESULT 17

US-09-949-016-12754  
; Sequence 12754, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12754  
; LENGTH: 29771  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(29771)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-12754

## Alignment Scores:

Pred. No.: 4.55e+04  
Score: 27.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
Indels: 0  
Gaps: 0

US-10-014-101B-41 (1-6) x US-09-949-016-12754 (1-29771)

QY 1 IleThrArgAlaArgile 6

Db 1375 ATAACCTAGGCCAGGATA 1392

## RESULT 18

US-09-949-016-13956  
; Sequence 13956, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13956  
; LENGTH: 29771  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(29771)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-13956

## Alignment Scores:

Pred. No.: 4.55e+04  
Score: 27.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
Indels: 0  
Gaps: 0

US-10-014-101B-41 (1-6) x US-09-949-016-13956 (1-29771)

QY 1 IleThrArgAlaArgile 6

Db 1375 ATAACCTAGGCCAGGATA 1392

## RESULT 19

```
US-09-949-016-17110
; Sequence 17110, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17110
; LENGTH: 70014
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(70014)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17110

Alignment Scores:
Pred. No.: 1.09e+05 Length: 70014
Score: 27.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-014-101B-41 (1-6) x US-09-949-016-17110 (1-70014)

QY 1 IleThrArgAlaArgIle 6
|||||
Db 31410 ATAACTAGGCGAAGATT 31427

RESULT 20
US-09-949-016-17490/c
; Sequence 17490, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17490
; LENGTH: 14115
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(14115)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17490

Alignment Scores:
Pred. No.: 2.21e+05 Length: 14115
Score: 161 Matches: 5
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 83.33% Mismatches: 0
Query Match: 96.30% Indels: 0
DB: 3 Gaps: 0

US-10-014-101B-41 (1-6) x US-09-949-016-17490 (1-14115)

QY 1 IleThrArgAlaArgIle 6
|||||
Db 59860 ATCACAGAGCCCGGATC 59843

RESULT 21
US-08-671-824-20/c
; Sequence 20, Application US/08671824
; Patent No. 6025167
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Zaug, Arthur J.
; APPLICANT: Been, Michael D.
; TITLE OF INVENTION: RNA RIBOZYME POLYMERASES,
; DEPHOSPHORYLASES, RESTRICTION
; ENDOURIBONUCLEASES AND METHODS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/671,824
; FILING DATE: June 5, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/278,624
; FILING DATE: July 21, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 220/166
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 75 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-671-824-20

Alignment Scores:
Pred. No.: 161 Length: 75
Score: 26.00 Matches: 5
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 83.33% Mismatches: 0
Query Match: 96.30% Indels: 0
DB: 3 Gaps: 0

US-10-014-101B-41 (1-6) x US-08-671-824-20 (1-75)

QY 1 IleThrArgAlaArgIle 6
|||||
```

```
Db      33 GTTACCGAGCTCGAATT 16
RESULT 22
US-09-686-341-16/c
; Sequence 16, Application US/09686341
; Patent No. 6696250
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Zaug, Arthur J.
; APPLICANT: Been, Michael D.
; TITLE OF INVENTION: RNA RIBOZYME POLYMERASES, DEPHOSPHORYLASES, RESTRICTION
; FILE REFERENCE: ENDOBONUCLEASES AND METHODS
; CURRENT APPLICATION NUMBER: US/09/686,341
; CURRENT FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: 09/005,325
; PRIOR FILING DATE: 1998-01-09
; PRIOR APPLICATION NUMBER: 08/671,824
; PRIOR FILING DATE: 1996-06-05
; PRIOR APPLICATION NUMBER: 08/278,624
; PRIOR FILING DATE: 1994-07-21
; PRIOR APPLICATION NUMBER: 07/843,737
; PRIOR FILING DATE: 1992-02-28
; PRIOR APPLICATION NUMBER: 07/562,672
; PRIOR FILING DATE: 1990-08-03
; PRIOR APPLICATION NUMBER: 06/937,327
; PRIOR FILING DATE: 1986-12-03
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 75
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthesized
; OTHER INFORMATION: nucleic acid molecule.
US-09-686-341-16

Alignment Scores:
Pred. No.: 161      Length: 75
Score: 26.00      Matches: 5
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 83.33%      Mismatches: 0
Query Match: 96.30%      Indels: 0
DB: 4      Gaps: 0

US-10-014-101B-41 (1-6) x US-09-686-341-16 (1-75)

Qy      1 IleThrArgAlaArgile 6
Db      33 GTTACCGAGCTCGAATT 16

RESULT 23
US-09-313-294A-2688
; Sequence 2688, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 2688
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700553119H1

Alignment Scores:
Pred. No.: 161      Length: 185
Score: 26.00      Matches: 5
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 83.33%      Mismatches: 0
Query Match: 96.30%      Indels: 0
DB: 4      Gaps: 0

US-10-014-101B-41 (1-6) x US-09-686-341-16 (1-75)

Qy      1 IleThrArgAlaArgile 6
Db      33 GTTACCGAGCTCGAATT 16

RESULT 24
US-09-313-294A-3534/c
; Sequence 3534, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 3534
; LENGTH: 283
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700611809H1

Alignment Scores:
Pred. No.: 640      Length: 283
Score: 26.00      Matches: 5
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 83.33%      Mismatches: 0
Query Match: 96.30%      Indels: 0
DB: 4      Gaps: 0

US-10-014-101B-41 (1-6) x US-09-313-294A-3534 (1-283)

Qy      1 IleThrArgAlaArgile 6
Db      23 ATCACCCGTCGCGAGTG 6

RESULT 25
US-09-313-294A-7329
; Sequence 7329, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 7329
; LENGTH: 291
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700553119H1
```

```
; NAME/KEY: unsure
; LOCATION: 24, 150
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-2688

Alignment Scores:
Pred. No.: 411      Length: 185
Score: 26.00      Matches: 5
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 83.33%      Mismatches: 0
Query Match: 96.30%      Indels: 0
DB: 4      Gaps: 0

US-10-014-101B-41 (1-6) x US-09-313-294A-2688 (1-185)

Qy      1 IleThrArgAlaArgile 6
Db      55 ATCACCTAGACGCGCTA 72

RESULT 24
US-09-313-294A-3534/c
; Sequence 3534, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 3534
; LENGTH: 283
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700611809H1

Alignment Scores:
Pred. No.: 640      Length: 283
Score: 26.00      Matches: 5
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 83.33%      Mismatches: 0
Query Match: 96.30%      Indels: 0
DB: 4      Gaps: 0

US-10-014-101B-41 (1-6) x US-09-313-294A-3534 (1-283)

Qy      1 IleThrArgAlaArgile 6
Db      23 ATCACCCGTCGCGAGTG 6

RESULT 25
US-09-313-294A-7329
; Sequence 7329, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 7329
; LENGTH: 291
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700553119H1
```

```
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6476212 700381547H1
; LOCATION: 4, 23
; NAME/KEY: unsure
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-7329

Alignment Scores:
Pred. No.: 659 Length: 291
Score: 26.00 Matches: 5
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 83.33% Mismatches: 0
Query Match: 96.30% Indels: 0
DB: 4 Gaps: 0

US-10-014-101B-41 (1-6) x US-09-313-294A-7329 (1-291)
QY 1 IleThrArgAlaArgIle 6
Db 12 ATACGAGACGAGGTT 29

RESULT 26
US-09-313-294A-4423
; Sequence 4423, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Laigudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 4423
; LENGTH: 297
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6476212 700348396H1
; NAME/KEY: unsure
; LOCATION: 195, 215, 243, 257, 264-265, 269, 280
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-4423

Alignment Scores:
Pred. No.: 673 Length: 297
Score: 26.00 Matches: 5
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 83.33% Mismatches: 0
Query Match: 96.30% Indels: 0
DB: 4 Gaps: 0

US-10-014-101B-41 (1-6) x US-09-313-294A-4423 (1-297)
QY 1 IleThrArgAlaArgIle 6
Db 27 ATACGAGACGAGGTT 44

RESULT 27
US-09-270-767-5082/c
; Sequence 5082, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
```

```
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5082
; LENGTH: 473
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-5082

Alignment Scores:
Pred. No.: 1,09e+03 Length: 473
Score: 26.00 Matches: 5
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 83.33% Mismatches: 0
Query Match: 96.30% Indels: 0
DB: 4 Gaps: 0

US-10-014-101B-41 (1-6) x US-09-270-767-5082 (1-473)
QY 1 IleThrArgAlaArgIle 6
Db 301 GTCACCTCGAGCCAGGATT 284

RESULT 28
US-09-270-767-20364/c
; Sequence 20364, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20364
; LENGTH: 473
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-20364

Alignment Scores:
Pred. No.: 1,09e+03 Length: 473
Score: 26.00 Matches: 5
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 83.33% Mismatches: 0
Query Match: 96.30% Indels: 0
DB: 4 Gaps: 0

US-10-014-101B-41 (1-6) x US-09-270-767-20364 (1-473)
QY 1 IleThrArgAlaArgIle 6
Db 301 GTCACCTCGAGCCAGGATT 284

RESULT 29
US-09-949-016-33777
; Sequence 33777, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
```

; SEQ ID NO 33777  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-33777

Alignment Scores:  
Pred. No.: 1.4e+03 Length: 601  
Score: 26.00 Matches: 5  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 83.33% Mismatches: 0  
Query Match: 96.30% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-41 (1-6) x US-09-949-016-33777 (1-601)

QY 1 IleThrArgAlaArgIle 6  
Db 83 ATCAAGGGCCCGTGTG 100

## RESULT 30

US-09-949-016-44357/c

; Sequence 44357, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 44357

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-44357

Alignment Scores:

Pred. No.: 1.4e+03 Length: 601  
Score: 26.00 Matches: 5  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 83.33% Mismatches: 0  
Query Match: 96.30% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-41 (1-6) x US-09-949-016-44357 (1-601)

QY 1 IleThrArgAlaArgIle 6  
Db 418 ATTCAAGGGCCCGTAGTA 401

## RESULT 31

US-09-949-016-44358/c

; Sequence 44358, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 44358  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-44358

Alignment Scores:  
Pred. No.: 1.4e+03 Length: 601  
Score: 26.00 Matches: 5  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 83.33% Mismatches: 0  
Query Match: 96.30% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-41 (1-6) x US-09-949-016-44358 (1-601)

QY 1 IleThrArgAlaArgIle 6  
Db 401 ATTCAAGGGCTAGAGTA 384

## RESULT 32

US-09-949-016-52245

; Sequence 52245, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 52245

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-52245

Alignment Scores:

Pred. No.: 1.4e+03 Length: 601  
Score: 26.00 Matches: 5  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 83.33% Mismatches: 0  
Query Match: 96.30% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-41 (1-6) x US-09-949-016-52245 (1-601)

QY 1 IleThrArgAlaArgIle 6  
Db 354 GTGACTCGAGCCAGGATC 371

## RESULT 33

US-09-949-016-52246

; Sequence 52246, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

```
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 52246
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-52246
```

```
Alignment Scores:
Pred. No.: 1.4e+03 Length: 601
Score: 26.00 Matches: 5
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 83.33% Mismatches: 0
Query Match: 96.30% Indels: 0
DB: 4 Gaps: 0
```

US-10-014-101B-41 (1-6) x US-09-949-016-52246 (1-601)

```
QY 1 IleThrArgAlaArgile 6
   ::::::::::::::::::::
DB 188 GTGACTCGAGCCAGGATC 205
```

## RESULT 34

```
US-09-949-016-60727/c
; Sequence 60727, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
```

```
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60727
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-60727
```

```
Alignment Scores:
Pred. No.: 1.4e+03 Length: 601
Score: 26.00 Matches: 5
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 83.33% Mismatches: 0
Query Match: 96.30% Indels: 0
DB: 4 Gaps: 0
```

US-10-014-101B-41 (1-6) x US-09-949-016-60727 (1-601)

```
QY 1 IleThrArgAlaArgile 6
   ::::::::::::::::::::
DB 283 ATCAAGGGCTAGATGA 266
```

## RESULT 35

```
US-09-949-016-110816
; Sequence 110816, Application US/09949016
; Patent No. 6812339
```

```
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 110816
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-110816
```

```
Alignment Scores:
Pred. No.: 1.4e+03 Length: 601
Score: 26.00 Matches: 5
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 83.33% Mismatches: 0
Query Match: 96.30% Indels: 0
DB: 4 Gaps: 0
```

US-10-014-101B-41 (1-6) x US-09-949-016-110816 (1-601)

```
QY 1 IleThrArgAlaArgile 6
   ::::::::::::::::::::
DB 523 GTCAAGGGCTCGTATT 540
```

## RESULT 36

```
US-09-949-016-110977
; Sequence 110977, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
```

```
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 110977
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-110977
```

```
Alignment Scores:
Pred. No.: 1.4e+03 Length: 601
Score: 26.00 Matches: 5
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 83.33% Mismatches: 0
Query Match: 96.30% Indels: 0
DB: 4 Gaps: 0
```

US-10-014-101B-41 (1-6) x US-09-949-016-110977 (1-601)

```
QY 1 IleThrArgAlaArgile 6
   ::::::::::::::::::::
DB 523 GTCAAGGGCTCGTATT 540
```

```
RESULT 37
US-09-949-016-111138
; Sequence 111138, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 111138
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-111138

Alignment Scores:
Pred. No.: 1.4e+03 Length: 601
Score: 26.00 Matches: 5
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 83.33% Mismatches: 0
Query Match: 96.30% Indels: 0
DB: 4 Gaps: 0

US-10-014-101B-41 (1-6) x US-09-949-016-111138 (1-601)
QY 1 IleThrArgAlaArgIle 6
Db 523 GTCACAAGGGCTCGTATT 540

RESULT 38
US-09-949-016-111299
; Sequence 111299, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 111299
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-111299

Alignment Scores:
Pred. No.: 1.4e+03 Length: 601
Score: 26.00 Matches: 5
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 83.33% Mismatches: 0
Query Match: 96.30% Indels: 0
DB: 4 Gaps: 0

US-10-014-101B-41 (1-6) x US-09-949-016-111299 (1-601)
QY 1 IleThrArgAlaArgIle 6
Db 83 ATCAACAGGGCCCGTGTG 100

RESULT 40
US-08-998-416-867
; Sequence 867, Application US/08998416
; Patent No. 6239284
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebschung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

APPLICATION NUMBER: US/08/998,416  
FILING DATE: 24-DEC-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: CH 0016/97  
FILING DATE: 31-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Meigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8587  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 867:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 714 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:   
ORGANISM: PAG1545RP  
US-08-998-416-867

Alignment Scores:  
Pred. No.: 1.68e+03 Length: 714  
Score: 26.00 Matches: 5  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 83.33% Mismatches: 0  
Query Match: 96.30% Indels: 0  
DB: 3 Gaps: 0

US-10-014-101B-41 (1-6) x US-08-998-416-867 (1-714)

Oy 1 IleThrArgAlaArgIle 6  
Db 439 GTCACCGGGCGAGGATT 456

Search completed: February 18, 2005, 07:19:22  
Job time : 55.65 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 18, 2005, 05:26:14 ; Search time 282.84 Seconds  
(without alignments)  
125.382 Million cell updates/sec

Title: US-10-014-101B-41

Perfect score: 27

Sequence: 1 ITRARI 6

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 5384158 seqs, 2955248155 residues

Total number of hits satisfying chosen parameters: 10768316

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10014101/runat\_16022005\_075848\_8061/app\_query.fasta\_1.796  
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb-MINMATCH=0.1  
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=40 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10014101 @CGN 1 1 659 @runat\_16022005\_075848\_8061  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq:\*
- 19: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*
- 20: /cgn2\_6/ptodata/2/pubpna/US11\_PUB.seq:\*
- 21: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 22: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	27	100.0	174	18	US-10-425-115-9138	Sequence 9138, Ap
C 2	27	100.0	175	18	US-10-437-963-10568	Sequence 10568, A
C 3	27	100.0	217	17	US-10-424-599-106272	Sequence 106272, A
C 4	27	100.0	289	18	US-10-425-115-23471	Sequence 23471, A
C 5	27	100.0	351	18	US-10-425-115-91752	Sequence 91752, A
C 6	27	100.0	369	18	US-10-425-115-74928	Sequence 74928, A
C 7	27	100.0	375	17	US-10-276-774-1129	Sequence 1129, Ap
C 8	27	100.0	384	9	US-09-983-965-3288	Sequence 3288, Ap
C 9	27	100.0	393	17	US-10-424-599-86559	Sequence 86559, A
C 10	27	100.0	406	18	US-10-437-963-34746	Sequence 34746, A
C 11	27	100.0	408	18	US-10-425-115-11568	Sequence 11568, A
C 12	27	100.0	463	18	US-10-425-115-110170	Sequence 110170, A
C 13	27	100.0	493	18	US-10-357-930-7957	Sequence 7957, Ap
C 14	27	100.0	524	17	US-10-424-599-32616	Sequence 32616, A
C 15	27	100.0	537	18	US-10-363-345A-10705	Sequence 10705, A
C 16	27	100.0	537	18	US-10-363-345A-10706	Sequence 10706, A
C 17	27	100.0	588	18	US-10-425-115-94827	Sequence 94827, A
C 18	27	100.0	624	18	US-10-437-963-18174	Sequence 18174, A
C 19	27	100.0	685	13	US-10-027-632-23732	Sequence 23732, A
C 20	27	100.0	685	17	US-10-027-632-23732	Sequence 23732, A
C 21	27	100.0	694	17	US-10-425-114-33876	Sequence 33876, A
C 22	27	100.0	694	18	US-10-425-115-35576	Sequence 35576, A
C 23	27	100.0	708	17	US-10-282-122A-9537	Sequence 9537, Ap
C 24	27	100.0	721	17	US-10-425-114-2055	Sequence 2055, Ap
C 25	27	100.0	737	13	US-10-027-632-21005	Sequence 21005, A
C 26	27	100.0	737	17	US-10-027-632-21005	Sequence 21005, A
C 27	27	100.0	750	15	US-10-156-761-5459	Sequence 5459, Ap
C 28	27	100.0	755	18	US-10-425-115-183342	Sequence 183342, A
C 29	27	100.0	807	17	US-10-425-114-15687	Sequence 15687, A
C 30	27	100.0	808	17	US-10-425-114-14429	Sequence 14429, A
C 31	27	100.0	840	18	US-10-425-115-35579	Sequence 35579, A
C 32	27	100.0	869	18	US-10-425-115-65670	Sequence 65670, A
C 33	27	100.0	890	18	US-10-425-115-51212	Sequence 51212, A
C 34	27	100.0	916	18	US-10-437-963-100905	Sequence 100905, A
C 35	27	100.0	955	18	US-10-425-115-35581	Sequence 35581, A
C 36	27	100.0	975	18	US-10-437-963-63753	Sequence 63753, A
C 37	27	100.0	996	18	US-10-425-115-65672	Sequence 65672, A
C 38	27	100.0	1023	18	US-10-425-115-65676	Sequence 65676, A
C 39	27	100.0	1023	17	US-10-424-599-87248	Sequence 87248, A
C 40	27	100.0	1033	17	US-10-425-114-33335	Sequence 33335, A
C 41	27	100.0	1039	18	US-10-357-930-29255	Sequence 29255, A
C 42	27	100.0	1043	18	US-10-425-115-65669	Sequence 65669, A
C 43	27	100.0	1059	15	US-10-146-772-297	Sequence 297, App
C 44	27	100.0	1059	17	US-10-241-742-297	Sequence 297, App
C 45	27	100.0	1059	17	US-10-440-523-297	Sequence 297, App

# ALIGNMENTS

RESULT 1  
US-10-425-115-9138/c  
; Sequence 9138, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(5322)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 9138  
; LENGTH: 174  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:

AFFILIATION: CAO YONGWEI  
;  
; ARSIZCANIZ:  
; ZHANG  
; APPLICANT: Cao, Yongwei

```

Qy      1 lleThrArgAlaArgIle 6
        |||||
Db      330 ATAACCGAGCTCGGATA 313

RESULT 9
US-10-424-599-86559
; Sequence 86559, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalick David K
; APPLICANT: Zhou Yihua

```

; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 86559  
; LENGTH: 393  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_49174C.1  
US-10-424-599-86559

Alignment Scores:  
Pred. No.: 1.59e+03 Length: 393  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 17 Gaps: 0

US-10-014-101B-41 (1-6) x US-10-424-599-86559 (1-393)

QY 1 IleThrArgAlaArgile 6  
Db 92 ATTACACGGCGCAAGGATC 109

## RESULT 10

US-10-437-963-34746/c

; Sequence 34746, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 34746

; LENGTH: 406

; TYPE: DNA

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_38731C.1

US-10-437-963-34746

Alignment Scores:  
Pred. No.: 1.64e+03 Length: 406  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 18 Gaps: 0

US-10-014-101B-41 (1-6) x US-10-437-963-34746 (1-406)

QY 1 IleThrArgAlaArgile 6  
Db 239 ATAACTCGGCGCGGCATC 222

## RESULT 11

US-10-425-115-111568/c

; Sequence 111568, Application US/10425115

; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 111568  
; LENGTH: 408  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_33238C.1  
US-10-425-115-111568

Alignment Scores:  
Pred. No.: 1.65e+03 Length: 408  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 18 Gaps: 0

US-10-014-101B-41 (1-6) x US-10-425-115-111568 (1-408)

QY 1 IleThrArgAlaArgile 6  
Db 21 ATTACACGGCGCGCATC 4

## RESULT 12

US-10-425-115-110170/c

; Sequence 110170, Application US/10425115

; Publication No. US20040214272A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53222)B

; CURRENT APPLICATION NUMBER: US/10/425,115

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 110170

; LENGTH: 463

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: MRT4577\_31970C.1

US-10-425-115-110170

Alignment Scores:  
Pred. No.: 1.87e+03 Length: 463  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 18 Gaps: 0

US-10-014-101B-41 (1-6) x US-10-425-115-110170 (1-463)

QY 1 IleThrArgAlaArgile 6  
Db 21 ATTACACGTGCACGCATC 4

## RESULT 13

US-10-357-930-7957

```
; Sequence 7957, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Engede, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7957
; LENGTH: 493
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 64, 68, 81, 83, 85, 96, 133, 147, 155, 159, 178, 190, 194,
; LOCATION: 195, 210, 222, 224, 225, 230, 242, 255, 270, 275, 281, 289,
; LOCATION: 291, 297, 309, 310, 315, 318, 324, 335, 339, 352, 353, 357,
; LOCATION: 362, 398, 400, 401, 408, 413, 420, 425, 427, 431, 434
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 439, 453, 455, 461, 466, 482, 484
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-7957

Alignment Scores:
Pred. No.: 1.98e+03 Length: 493
Score: 27.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-10-014-101B-41 (1-6) x US-10-357-930-7957 (1-493)

QY 1 IleThrArgAlaArgile 6
Db 98 ATTACCGTCGCCGAGATA 115

RESULT 14
US-10-424-599-32616
; Sequence 32616, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 32616
; LENGTH: 524
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_129457C.1
US-10-424-599-32616

Alignment Scores:
Pred. No.: 2.1e+03 Length: 524
Score: 27.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-10-014-101B-41 (1-6) x US-10-424-599-32616 (1-524)

QY 1 IleThrArgAlaArgile 6
Db 457 ATCACAGCAGCCAGGATA 474

RESULT 15
US-10-363-345A-10705/c
; Sequence 10705, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; TITLE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3
; FILE REFERENCE: E01/1227
; CURRENT APPLICATION NUMBER: US/10/363,345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 10705
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 10705
US-10-363-345A-10705

Alignment Scores:
Pred. No.: 2.15e+03 Length: 537
Score: 27.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-10-014-101B-41 (1-6) x US-10-363-345A-10705 (1-537)

QY 1 IleThrArgAlaArgile 6
Db 298 ATAACCCGCGCAGGATA 281

RESULT 16
US-10-363-345A-10706
; Sequence 10706, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; TITLE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3
; FILE REFERENCE: E01/1227
; CURRENT APPLICATION NUMBER: US/10/363,345A
; CURRENT FILING DATE: 2003-03-03
```

; NUMBER OF SEQ ID NOS: 40712  
; SEQ ID NO 10706  
; LENGTH: 537  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
; OTHER INFORMATION: CpG-island No: 10706  
US-10-363-345A-10706

Alignment Scores:  
Pred. No.: 2.15e+03 Length: 537  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 18 Gaps: 0

US-10-014-101B-41 (1-6) x US-10-363-345A-10706 (1-537)

QY 1 IleThrArgAlaArgile 6  
|||  
DB 240 ATACCCGCGCAGATA 257

## RESULT 17

US-10-425-115-94827  
; Sequence 94827, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53222)B

; CURRENT APPLICATION NUMBER: US/10/425,115

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 94827

; LENGTH: 588

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: MRT4577\_186474C.1

US-10-425-115-94827

Alignment Scores:  
Pred. No.: 2.35e+03 Length: 588  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 18 Gaps: 0

US-10-014-101B-41 (1-6) x US-10-425-115-94827 (1-588)

QY 1 IleThrArgAlaArgile 6  
|||  
DB 101 ATACCCGCGCCCGATC 118

## RESULT 18

US-10-437-963-18174  
; Sequence 18174, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 18174  
; LENGTH: 624  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_23758C.1  
US-10-437-963-18174

Alignment Scores:  
Pred. No.: 2.49e+03 Length: 624  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 18 Gaps: 0

US-10-014-101B-41 (1-6) x US-10-437-963-18174 (1-624)

QY 1 IleThrArgAlaArgile 6  
|||  
DB 86 ATTACGCGCGCAAGATC 103

## RESULT 19

US-10-027-632-23732/c  
; Sequence 23732, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632

; CURRENT FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/198,676

; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/193,483

; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23

; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002

; PRIOR FILING DATE: 1999-08-09

; NUMBER OF SEQ ID NOS: 325720

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 23732

; LENGTH: 685

; TYPE: DNA

; ORGANISM: Human

US-10-027-632-23732

Alignment Scores:  
Pred. No.: 2.72e+03 Length: 685  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 13 Gaps: 0

US-10-014-101B-41 (1-6) x US-10-027-632-23732 (1-685)

QY 1 IleThrArgAlaArgile 6

```
Db      445 ATTACAGGGCTAGAATT 428
|||||
RESULT 20
US-10-027-632-23732/c
; Sequence 23732, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;   POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/195,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23732
;   LENGTH: 685
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-23732
Alignment Scores:
Pred. No.:      2,72e+03      Length:      685
Score:          27.00         Matches:      6
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches:      0
Query Match:    100.00%      Indels:        0
DB:             17           Gaps:         0

US-10-014-101B-41 (1-6) x US-10-027-632-23732 (1-685)

QY      1 IleThrArgAlaArgIle 6
|||||
Db      445 ATTACAGGGCTAGAATT 428
|||||
RESULT 21
US-10-425-114-33876
; Sequence 33876, Application US/10425114
; Publication No. US20040034898A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;   TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 33876
;   LENGTH: 694
;   TYPE: DNA
;   ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17179H04_FLI

US-10-014-101B-41 (1-6) x US-10-425-114-33876 (1-694)

QY      1 IleThrArgAlaArgIle 6
|||||
Db      262 ATAACCCAGCAAGGATT 279
|||||
RESULT 22
US-10-425-115-35576
; Sequence 35576, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;   TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 35576
;   LENGTH: 694
;   TYPE: DNA
;   ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_132443C.1

US-10-425-115-35576
Alignment Scores:
Pred. No.:      2,76e+03      Length:      694
Score:          27.00         Matches:      6
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches:      0
Query Match:    100.00%      Indels:        0
DB:             18           Gaps:         0

US-10-014-101B-41 (1-6) x US-10-425-115-35576 (1-694)

QY      1 IleThrArgAlaArgIle 6
|||||
Db      262 ATAACCCAGCAAGGATT 279
|||||
RESULT 23
US-10-282-122A-9537
; Sequence 9537, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
```

; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 9537  
; LENGTH: 708  
; TYPE: DNA  
; ORGANISM: Bacillus anthracis  
US-10-282-122A-9537

Alignment Scores:  
Pred. No.: 2,81e+03 Length: 708  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 17 Gaps: 0

US-10-014-101B-41 (1-6) x US-10-282-122A-9537 (1-708)

QY 1 IleThrArgAlaArgile 6  
Db 150 ATTACAGGCGGAGGATA 167

RESULT 24  
US-10-425-114-2055  
; Sequence 2055, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 2055  
; LENGTH: 721  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700201559\_FLI  
US-10-425-114-2055

Alignment Scores:  
Pred. No.: 2,86e+03 Length: 721  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 17 Gaps: 0  
US-10-014-101B-41 (1-6) x US-10-425-114-2055 (1-721)

QY 1 IleThrArgAlaArgile 6  
Db 598 ATTACAGGCGGAGGATT 615

RESULT 25  
US-10-027-632-21005  
; Sequence 21005, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 21005  
; LENGTH: 737  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-21005

Alignment Scores:  
Pred. No.: 2,92e+03 Length: 737  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 13 Gaps: 0

US-10-014-101B-41 (1-6) x US-10-027-632-21005 (1-737)

QY 1 IleThrArgAlaArgile 6  
Db 525 ATTACAGGCGGAGGATT 542

RESULT 26  
US-10-027-632-21005  
; Sequence 21005, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483

; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 21005  
; LENGTH: 737  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-21005

Alignment Scores: Length: 737  
Pred. No.: 2.92e+03 Matches: 6  
Score: 27.00 Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 100.00% Gaps: 0  
DB: 17

US-10-014-101B-41 (1-6) x US-10-027-632-21005 (1-737)

Oy 1 IleThrArgAlaArgile 6  
|||||  
Db 525 ATTACAGGCGTAGAATT 542

## RESULT 27

US-10-156-761-5459/c  
; Sequence 5459, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; PRIOR FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 5459  
; LENGTH: 750  
; TYPE: DNA  
; ORGANISM: Streptomyces avermitilis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(750)  
US-10-156-761-5459

Alignment Scores: Length: 750  
Pred. No.: 2.97e+03 Matches: 6  
Score: 27.00 Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 100.00% Gaps: 0  
DB: 15

US-10-014-101B-41 (1-6) x US-10-156-761-5459 (1-750)

Oy 1 IleThrArgAlaArgile 6  
|||||

Db 551 ATCAGAGGCGGAGAAATC 534

## RESULT 28

US-10-425-115-183342  
; Sequence 183342, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 183342  
; LENGTH: 755  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_98793C.1  
US-10-425-115-183342

Alignment Scores: Length: 755  
Pred. No.: 2.99e+03 Matches: 6  
Score: 27.00 Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 100.00% Gaps: 0  
DB: 18

US-10-014-101B-41 (1-6) x US-10-425-115-183342 (1-755)

Oy 1 IleThrArgAlaArgile 6  
|||||  
Db 194 ATTACAGTGCACGCATT 211

## RESULT 29

US-10-425-114-15687  
; Sequence 15687, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 15687  
; LENGTH: 807  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3060-052-F10\_FLI  
US-10-425-114-15687

Alignment Scores: Length: 807  
Pred. No.: 3.19e+03 Matches: 6  
Score: 27.00 Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 100.00% Gaps: 0  
DB: 17

US-10-014-101B-41 (1-6) x US-10-425-114-15687 (1-807)

```
QY      1 ileThrArgAlaArgile 6
Db      681 ATTACACGTGCACGCATT 698

RESULT 30
US-10-425-114-14429
; Sequence 14429, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 14429
; LENGTH: 808
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB189-022-H10_FLI
US-10-425-114-14429

Alignment Scores:
Pred. No.: 3.19e+03      Length: 808
Score: 27.00             Matches: 6
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 100.00%             Indels: 0
DB: 17                   Gaps: 0

US-10-014-101B-41 (1-6) x US-10-425-114-14429 (1-808)

QY      1 ileThrArgAlaArgile 6
Db      684 ATTACACGTGCACGCATT 701

RESULT 31
US-10-425-115-35579
; Sequence 35579, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 35579
; LENGTH: 840
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(840)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_132446C.1
US-10-425-115-35579

Alignment Scores:
Pred. No.: 3.32e+03      Length: 840
Score: 27.00             Matches: 6
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 100.00%             Indels: 0
DB: 17                   Gaps: 0

US-10-014-101B-41 (1-6) x US-10-425-115-35579 (1-840)

QY      1 ileThrArgAlaArgile 6
Db      677 ATCACCAGCAAGCATT 694

RESULT 32
US-10-425-115-65670
; Sequence 65670, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 65670
; LENGTH: 869
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_159897C.1
US-10-425-115-65670

Alignment Scores:
Pred. No.: 3.43e+03      Length: 869
Score: 27.00             Matches: 6
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 100.00%             Indels: 0
DB: 18                   Gaps: 0

US-10-014-101B-41 (1-6) x US-10-425-115-65670 (1-869)

QY      1 ileThrArgAlaArgile 6
Db      700 ATTACACGTGCACGCATT 717

RESULT 33
US-10-425-115-51212
; Sequence 51212, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 51212
; LENGTH: 890
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_146701C.1
US-10-425-115-51212

Alignment Scores:
Pred. No.: 3.43e+03      Length: 890
Score: 27.00             Matches: 6
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 100.00%             Indels: 0
DB: 18                   Gaps: 0
```

Pred. No.: 3.51e+03 Length: 890  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 0 Gaps: 0

US-10-014-101B-41 (1-6) x US-10-425-115-51212 (1-890)

Qy 1 IleThrArgAlaArgIle 6  
Db 70 ATTACAGAGCTCGCATA 87

## RESULT 34

US-10-437-963-100905  
; Sequence 100905, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 100905  
; LENGTH: 916  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clones ID: PAT\_MRT4530\_98576C.1  
US-10-437-963-100905

Alignment Scores:  
Pred. No.: 3.61e+03 Length: 916  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 0 Gaps: 0

US-10-014-101B-41 (1-6) x US-10-437-963-100905 (1-916)

Qy 1 IleThrArgAlaArgIle 6  
Db 809 ATAACAAGGCCAGGATT 826

## RESULT 35

US-10-425-115-35581  
; Sequence 35581, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 35581  
; LENGTH: 955  
; TYPE: DNA  
; ORGANISM: Zea mays

; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(955)  
; OTHER INFORMATION: unsure at all n locations  
; FEATURE:  
; OTHER INFORMATION: Clones ID: MRT4577\_132448C.1  
US-10-425-115-35581

Alignment Scores:  
Pred. No.: 3.75e+03 Length: 955  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 0 Gaps: 0

US-10-014-101B-41 (1-6) x US-10-425-115-35581 (1-955)

Qy 1 IleThrArgAlaArgIle 6  
Db 826 ATCACCAGAGCAAGGATT 843

## RESULT 36

US-10-437-963-63753  
; Sequence 63753, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 63753  
; LENGTH: 975  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clones ID: PAT\_MRT4530\_64965C.1  
US-10-437-963-63753

Alignment Scores:  
Pred. No.: 3.83e+03 Length: 975  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 0 Gaps: 0

US-10-014-101B-41 (1-6) x US-10-437-963-63753 (1-975)

Qy 1 IleThrArgAlaArgIle 6  
Db 62 ATAACAAGAGCTCGATA 79

## RESULT 37

US-10-425-115-65672  
; Sequence 65672, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

;  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 65672  
; LENGTH: 996  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_159899C.1  
US-10-425-115-65672

Alignment Scores:  
Pred. No.: 3 91e+03 Length: 996  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 18 Gaps: 0

US-10-014-101B-41 (1-6) x US-10-425-115-65672 (1-996)

QY 1 IleThrArgAlaArgile 6  
|||||  
DB 840 ATTACACGTGCACGCATC 857

## RESULT 38

US-10-425-115-65676  
; Sequence 65676, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 65676  
; LENGTH: 1023  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_159901C.1  
US-10-425-115-65676

Alignment Scores:  
Pred. No.: 4 01e+03 Length: 1023  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 18 Gaps: 0

US-10-014-101B-41 (1-6) x US-10-425-115-65676 (1-1023)

QY 1 IleThrArgAlaArgile 6  
|||||  
DB 854 ATTACACGTGCACGCATT 871

## RESULT 39

US-10-424-599-87248/c  
; Sequence 87248, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei

;  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 87248  
; LENGTH: 1027  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_49796C.1  
US-10-424-599-87248

Alignment Scores:  
Pred. No.: 4 03e+03 Length: 1027  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 17 Gaps: 0

US-10-014-101B-41 (1-6) x US-10-424-599-87248 (1-1027)

QY 1 IleThrArgAlaArgile 6  
|||||  
DB 222 ATACACGCGCGAGGATA 205

## RESULT 40

US-10-425-114-32335  
; Sequence 32335, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 32335  
; LENGTH: 1033  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-ZMFLB73297H03\_FLI  
US-10-425-114-32335

Alignment Scores:  
Pred. No.: 4 05e+03 Length: 1033  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 17 Gaps: 0

US-10-014-101B-41 (1-6) x US-10-425-114-32335 (1-1033)

QY 1 IleThrArgAlaArgile 6  
|||||  
DB 24 ATACTCGCGCCGATC 41

Search completed: February 18, 2005, 10:15:51  
Job time : 284.84 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 18, 2005, 02:20:24 ; Search time 707.52 Seconds  
(without alignments)  
322.797 Million cell updates/sec

Title: US-10-014-101B-41

Perfect score: 27

Sequence: 1 ITRARI 6

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5	
Fgapop 6.0	Fgapext 7.0	
Delop 6.0	Delext 7.0	

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q/cg2\_1/USPTO.spool/US10014101/runat 16022005 075846 7952/app query.fasta\_1.796  
-DB=EST -QPM=fastap -SUFFIX=rat -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10014101 @CGN 1 1 5453 @runat 16022005 075846 7952 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gsal:\*  
9: gb\_g982:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	27	100.0	67	9 BX536548	Arabidops
C 2	27	100.0	100	2 BF174648	WHE3555A
C 3	27	100.0	121	4 BG068639	H3067G04
C 4	27	100.0	136	7 CK104188	I018P53.5
C 5	27	100.0	140	6 CD062932	MA1-0028U
C 6	27	100.0	152	2 AW287771	829007E10
C 7	27	100.0	161	8 BZ637306	OSCBF79TC
C 8	27	100.0	171	8 AZ023516	RPCI-23-3
C 9	27	100.0	179	6 CB280935	Jaa05d01.

C 10	27	100.0	183	9 CT98738	SALK_1473
C 11	27	100.0	186	6 CA485620	WHE4321_A
C 12	27	100.0	194	6 CA452616	Kt1N-4_F0
C 13	27	100.0	195	6 CA452634	Kt1N-2_B0
C 14	27	100.0	214	4 BG412344	OV2_40_G0
C 15	27	100.0	224	8 BH868282	hk23h11_x
C 16	27	100.0	233	2 BB369534	BB369534
C 17	27	100.0	235	2 BE764604	PM2-NT007
C 18	27	100.0	237	4 B1190218	h1f02E8.f
C 19	27	100.0	239	4 BG994575	RC4-HT114
C 20	27	100.0	249	2 BE025443	945028H10
C 21	27	100.0	253	6 CD712578	VVB021G12
C 22	27	100.0	258	4 BM078603	MEST122-A
C 23	27	100.0	259	1 AV323709	AV323709
C 24	27	100.0	260	2 AW620066	819_MARC
C 25	27	100.0	261	4 B1055989	PM2-GN043
C 26	27	100.0	261	6 CA811112	CA22LI041
C 27	27	100.0	262	8 BB319596	BB319596
C 28	27	100.0	263	8 BB1555	CIT-HSP-206
C 29	27	100.0	265	7 CR516082	CR516082
C 30	27	100.0	266	2 BE529872	M75N15STM
C 31	27	100.0	266	7 CK691445	ZF101-P00
C 32	27	100.0	267	2 BB577396	BB577396
C 33	27	100.0	267	8 BZ671989	PUBK67TD
C 34	27	100.0	269	8 BH228153	1006144E1
C 35	27	100.0	279	6 CB921222	VVD068E05
C 36	27	100.0	285	6 CD152034	ML1-0027T
C 37	27	100.0	286	2 AW870159	NXNV_124
C 38	27	100.0	287	2 BB336254	BB336254
C 39	27	100.0	288	6 CD061127	MA1-0016U
C 40	27	100.0	289	4 BM348722	MEST239-G
C 41	27	100.0	290	9 BX198253	Danio rer
C 42	27	100.0	293	4 BM349587	MEST252-B
C 43	27	100.0	295	2 BE580170	kg38a03.Y
C 44	27	100.0	296	1 AU023430	AU023430
C 45	27	100.0	297	6 CB339189	CA23E1031

#### ALIGNMENTS

RESULT 1  
BX536548/c

LOCUS  
DEFINITION

BX536548 67 bp DNA linear GSS 04-APR-2004  
Arabidopsis thaliana T-DNA flanking sequence GK-528G02-020333,  
genomic survey sequence.

ACCESSION  
BX536548

VERSION  
BX536548.1 GI:31413678

KEYWORDS  
GSS

SOURCE  
Arabidopsis thaliana (thale cress)

ORGANISM  
Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi

REFERENCE

AUTHORS

TITLE

Li, Y., Rosso, M.G., Strizhov, N., Viehoveer, P. and Weissshaar, B.  
GABI-Kat SimpleSearch: a flanking sequence tag (FST) database for  
the identification of T-DNA insertion mutants in Arabidopsis  
thaliana

Bioinformatics 19 (11), 1441-1442 (2003)

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K. and  
Weissshaar, B.  
An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for  
flanking sequence tag-based reverse genetics

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

Strizhov, N., Li, Y., Rosso, M.G., Viehoveer, P., Dekker, K.A. and  
Weissshaar, B.  
High-throughput generation of sequence indexes from T-DNA

mutagenized Arabidopsis thaliana lines  
Biotechniques 35 (6), 1164-1168 (2003)

JOURNAL  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

4 (bases 1 to 67)  
Strizhov, N., Li, Y., Rosso, M.G. and Weisshaar, B.  
Direct Submission

Submitted (31-MAR-2004) Weisshaar B., Max-Planck-Institut fuer  
Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany  
This sequence has been recovered from the left border of the T-DNA.  
It indicates an insertion within the locus defined by BAC clone  
p9K20. Details on the protocols used for generation of the sequence  
are described in References 1-3. The sequences are generated at the  
MPI for Plant Breeding Research in the context of the GABI-Kat  
project. GABI-Kat is part of the German Plant Genomics program  
designated 'GABI'. Information on line availability can be found  
at: <http://www.mpiz-koeln.mpg.de/GABI-Kat/>.

## FEATURES

source

1..67  
Location/Qualifiers  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/strain="Columbia 0"  
/db\_xref="taxon:3702"  
/clone="GK-528G02-020333"  
/clone\_lib="Arabidopsis thaliana T-DNA insertion lines"  
/ecotype="Col-0"

/note="PCR was performed on DNA from Arabidopsis thaliana  
plants (Ti) which were transformed with the T-DNA from  
vector pAC161 (Genbank accession number: AJ537514). The  
lines contain one or more T-DNA insertions. The DNA  
fragment(s) resulting from the PCR were directly sequenced  
to determine the genomic sequence flanking the insertion.  
T-DNA derived sequences were removed."

## ORIGIN

Alignment Scores:  
Pred. No.: 1.51e+03 Length: 67  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-10-014-101B-41 (1-6) x BX536548 (1-67)

Qy 1 lIeThrArgAlaArgIle 6  
Db 26 ATAACACGAGCCAGAATT 9

RESULT 2  
BF174648 100 bp mRNA linear EST 23-MAR-2001  
LOCUS MYE3555a Myeloma (MYE) cDNA library Homo sapiens CDNA, mRNA  
DEFINITION

sequence.  
ACCESSION BF174648  
VERSION BF174648.1 GI:13440862  
KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 100)

Claudio, J.O., Masih-Khan, E., Tang, H., Goncalves, J., Voralia, M.,

Li, Z.H., Nadeem, V., Cukerman, E., Francisco-Pabalan, O., Liew, C.C.,

Woodgett, J.R. and Stewart, A.K.

A molecular compendium of genes expressed in multiple myeloma

Blood 100 (6), 2175-2186 (2002)

22188429

12200383

Contact: A. Keith Stewart, M.D.

Oncology Research

University Health Network

610 University Ave., 5-126, Toronto, Ontario, M5G 2M9, Canada

Tel: (416) 946-4639

Fax: (416) 946-6546

Email: [k.stewart@utoronto.ca](mailto:k.stewart@utoronto.ca)

PCR Primers

FORWARD: 5'-GCCAGCTCGAATTACCTCACTAAAGGG-3'

BACKWARD: 5'-CCAGTAATTGTAATAGCTACTATAGGGCG-3'

Seq primer: 5'-GAAATTAACCTCAAGG-3'

## FEATURES

source

1..100  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/sex="male"  
/tissue type="Blood"  
/cell type="myeloma"  
/dev stage="multiple myeloma"  
/clone\_lib="Myeloma (MYE) CDNA library"

/note="Vector: Lambda Zap Express; Site 1: EcoRI; Site 2:  
XhoI; Myeloma cells from multiple myeloma patients' bone  
marrow were purified by magnetic cell sorting. mRNA were  
purified and an oligo d(T)18 primer containing XhoI  
restriction site was used to prime first strand synthesis  
using M-MLV reverse transcriptase. To protect the cDNAs  
from XhoI digestion in subsequent cloning step, the  
nucleotide analogue 5-methyl-dCTP was added to the  
nucleotide mixture and la-32Pi dATP was added to monitor  
the quantity and quality of first strand synthesis. After  
second-strand synthesis and blunting of cDNA termini,  
EcoRI adapters were ligated, followed by kinase treatment  
and digestion with XhoI. The cDNAs were then  
size-fractionated using Sephadryl S-500 column and then  
ligated into EcoRI and XhoI digested Lambda Zap Express  
vector. The ligation product was packaged using Gigapack  
II packaging extract. The library had primary titre of  
approx. 1x10<sup>6</sup>. Clones from the primary library were  
randomly selected for single pass sequencing."

## ORIGIN

Alignment Scores:  
Pred. No.: 2.24e+03 Length: 100  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0

US-10-014-101B-41 (1-6) x BF174648 (1-100)

Qy 1 lIeThrArgAlaArgIle 6  
Db 65 ATAACCGAGCGAGGAATC 82

RESULT 3  
BG068639/c 121 bp mRNA linear EST 17-DEC-2003  
LOCUS H3067G04-3 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone  
DEFINITION

H3067G04 3', mRNA sequence.

ACCESSION BG068639

VERSION BG068639.2

KEYWORDS GI:40016430

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 121)

Tanaka, T.S., Jaradat, S.A., Lim, M.K., Kargul, G.J., Wang, X.,

Grubov, M.J., Pantano, S., Sano, Y., Piao, Y., Nagaraja, R., Doi, H.,

Wood, W.H. III, Becker, K.G. and Ko, M.S.H.

Genome-wide expression profiling of mid-gestation placenta and

embryo using a 15,000 mouse developmental cDNA microarray

Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)

20381348

10922068

On Jan 26, 2001 this sequence version replaced gi:12551208.

COMMENT

Other\_ESTs: H3067G04-5  
 Contact: George J. Kargul  
 Laboratory of Genetics  
 National Institute on Aging/National Institutes of Health  
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
 Email: cdnaelgaun.grc.nia.nih.gov  
 This clone set has been freely distributed to the community. Please  
 visit <http://igsun.grc.nia.nih.gov/cdna/15k.html> for details.  
 Plate: H3067 row: G column: 04  
 Seq primer: -21M13 Forward  
 High quality sequence stop: 121  
 POLYA=Yes.

#### FEATURES

source Location/Qualifiers  
 1..121  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="niaEST:H3067G04-3"  
 /db\_xref="taxon:10090"  
 /clone="H3067G04"  
 /sex="Clones arrayed from a variety of cDNA libraries"  
 /dev\_stage="Clones arrayed from a variety of cDNA libraries"  
 /lab\_host="DH10B"  
 /clone\_lib="NIA Mouse 15K cDNA Clone Set"  
 /notes="Vector: pSPOR1, Site 1: SalI; Site 2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U.S.A., 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

#### ORIGIN

Alignment Scores:  
 Pred. No.: 2..7e+03 Length: 121  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0

US-10-014-101B-41 (1-6) x BG068639 (1-121)

QY 1 IleThrArgAlaArgile 6  
 |||||  
 25 ATCAAGAGCCAGGATT 8

#### RESULT 4

CK104188/c  
 LOCUS CK104188 136 bp mRNA linear EST 01-DEC-2003  
 DEFINITION 1018P53-5Pr Populus senescing leaves cDNA library Populus tremula cDNA clone 1018P53 5', mRNA sequence.

ACCESSION CK104188  
 VERSION CK104188.1 GI:38588513  
 KEYWORDS EST.  
 SOURCE Populus tremula

#### ORGANISM

Populus tremula  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Malpighiales; Salicaceae; Populus.

#### REFERENCE

AUTHORS 1 (bases 1 to 136)  
 Brerky, F., Bhalarao, R.R., Unneberg, P., Segerman, B., Nilsson, P.,  
 Brunner, A.M., Campaa, L., Jonsson-Lindvall, J., Tandré, K.,  
 Strauss, S.H., Sundberg, B., Gustafsson, P., Uhlen, M., Bhalarao, R.P.,  
 Nilsson, O., Sandberg, G., Karlsson, J., Lundberg, J. and Jansson, S.  
 A Populus EST resource for functional genomics  
 Unpublished (2003)

#### JOURNAL

COMMENT Other\_ESTs: 1018P53P, 1018P53.3Pr  
 Contact: Bo Segerman  
 Umea Plant Science Center, Department of Plant Physiology  
 Umea University  
 901 87 Umea, Sweden  
 Tel: +46 90 786 5279  
 Fax: +46 90 786 6676  
 Email: bo.segerman@plantphys.umu.se.

#### FEATURES

source Location/Qualifiers  
 1..136  
 /organism="Populus tremula"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:113636"  
 /clone="1018P53"  
 /tissue\_type="Senescing leaves"  
 /clone\_lib="Populus senescing leaves cDNA library"

#### ORIGIN

Alignment Scores:  
 Pred. No.: 3..03e+03 Length: 136  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 7 Gaps: 0

US-10-014-101B-41 (1-6) x CK104188 (1-136)

QY 1 IleThrArgAlaArgile 6  
 |||||  
 109 ATCACCAGGCGAGCATT 92

#### RESULT 5

CD062932/c  
 LOCUS CD062932 140 bp mRNA linear EST 14-SEP-2003  
 DEFINITION MA1-0028U-A010-D07-U-B MA1-0028 Schistosoma mansoni cDNA clone  
 MA1-0028U-A010-D07.B, mRNA sequence.

ACCESSION CD062932  
 VERSION CD062932.1 GI:34614092

#### KEYWORDS

SOURCE Schistosoma mansoni  
 ORGANISM Schistosoma mansoni  
 Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;  
 Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.

#### REFERENCE

AUTHORS 1 (bases 1 to 140)  
 Verjovski-Almeida, S., DeMarco, R., Martins, E.A.L., Guimaraes, P.E.M.,  
 Ojopi, E.P.B., Paquola, A.C.M., Piazza, J.P., Nishiyama, M.Y. Jr.,  
 Kitajima, J.P., Adamson, R.E., Ashton, P.D., Bonaldo, M.F.,  
 Coulson, P.S., Dillon, G.P., Farías, L.P., Gregorio, S.P., Ho, P.L.,  
 Leite, R.A., Malaquias, L.C.C., Marques, R.C.P., Miyasato, P.A.,  
 Nascimento, A.L.T.O., Ohlweiler, F.P., Reis, E.M., Ribeiro, M.A.,  
 Sa, R.G., Stukart, G.C., Soares, M.B., Gargioni, C., Kawano, T.,  
 Rodrigues, V., Madeira, A.M.B.N., Wilson, R.A., Menck, C.F.M.,  
 Setubal, J.C., Leite, L.C.C. and Dias-Neto, E.  
 Transcriptome analysis of the acelomate human parasite Schistosoma  
 mansoni  
 Nat. Genet. 35 (2), 148-157 (2003)

#### JOURNAL

MEDLINE 22879926  
 PUBLISHED 12973350

#### COMMENT

Contact: Dr. Sergio Verjovski-Almeida  
 Departamento de Bioquímica  
 Instituto de Química - Universidade de São Paulo  
 Av. Prof. Lineu Prestes 748 sala 1200, 05508-900 São Paulo - SP,  
 Brasil  
 Tel: +55-11-3091-2173  
 Fax: +55-11-3091-2186

Email: verio@iq.usp.br  
 This sequence was derived from the FAPESP Schistosoma mansoni EST Genome Project. All sequences in the project were assembled and annotated. This entry and all the assembled sequences can be seen in the following URL <http://bioinfo.iq.usp.br/schisto/>  
 Plate: MA1-0028U-A010 row: 7 column: D.

#### FEATURES

Location/Qualifiers  
 1..140  
 /organism="Schistosoma mansoni"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:6183"  
 /clone="MA1-0028U-A010-D07.B"  
 /sex="mixed pool"  
 /dev\_stage="adult"  
 /lab\_host="Mesocricetus auratus"  
 /clone\_lib="MA1-0028"  
 /note="Vector: SureClone"

#### ORIGIN

Alignment Scores:  
 Pred. No.: 3.12e+03 Length: 140  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-10-014-101B-41 (1-6) x CD062932 (1-140)

Qy 1 IleThrArgAlaArgIle 6  
 |||||  
 Db 29 ATCACTCGTCGACGCATT 12

#### RESULT 6

AW287771/c 152 bp mRNA linear EST 09-FEB-2000  
 LOCUS 829007E10.x1 829 - Silk infected with Fusarium Zea mays cDNA, mRNA  
 DEFINITION sequence.

ACCESSION AW287771  
 VERSION AW287771.1 GI:6681784

KEYWORDS EST.

SOURCE Zea mays

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 152)

Walbot,V.

Maize ESTs from various cDNA libraries sequenced at Stanford University

Unpublished (1999)

Contact: Walbot V

Department of Biological Sciences

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221

Email: walbot@stanford.edu

Plate: 829007 row: E column: 10.

Location/Qualifiers

#### FEATURES

1..152  
 /organism="Zea mays"  
 /mol\_type="mRNA"  
 /cultivar="B73"  
 /db\_xref="taxon:4577"  
 /tissue\_type="silk"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"

/clone\_lib="829 - Silk infected with Fusarium"  
 /note="Organ: silk; Vector: pBluescript II XR; Site 1: xhoI; Site 2: EcoRI; cDNA library of silks infected with 1 microliter of 500,000 spores/ml solution of Fusarium graminearum DAOM 180378. Prepared by Sharon Allard of

#### ORIGIN

Alignment Scores:  
 Pred. No.: 3.38e+03 Length: 152  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0

US-10-014-101B-41 (1-6) x AW287771 (1-152)

Qy 1 IleThrArgAlaArgIle 6  
 |||||

Db 90 ATTACACGTGCACGCATT 73

#### RESULT 7

BZ637306/c 161 bp DNA linear GSS 29-JAN-2003  
 LOCUS OGCBF79TC ZM.0.7.1.5 KB Zea mays genomic clone ZMMBMA0132N14,  
 DEFINITION genomic survey sequence.

ACCESSION BZ637306

VERSION BZ637306.1 GI:28088139

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 161)

Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,

Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,

Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.

Consortium for Maize Genomics

Unpublished (2002)

Other\_GSSs: OGCBF79TM

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TF

Class: sheared ends.

Location/Qualifiers

source

1..161

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone\_lib="ZMMBMA0132N14"

/note="Vector: pBGSK; Site 1: HincII; 0.7-1.5 kb

methylation filtered genomic DNA library"

#### ORIGIN

Alignment Scores:  
 Pred. No.: 3.58e+03 Length: 161  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 8 Gaps: 0

US-10-014-101B-41 (1-6) x BZ637306 (1-161)

Qy 1 IleThrArgAlaArgIle 6  
 |||||

Db 35 ATACACGGGCAAGGATT 18

```

RESULT 8
AZ023516
LOCUS
DEFINITION
  RPCI-23-336P9.TV RPCI-23 Mus musculus genomic clone RPCI-23-336P9,
  genomic survey sequence.
ACCESSION
  AZ023516
VERSION
  AZ023516.1 GI:7098900
KEYWORDS
  GSS.
SOURCE
  Mus musculus (house mouse)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
  1 (bases 1 to 171)
AUTHORS
  Zhao,S., Niernan,W., Feldblyum,T., Malek,J., Shatsman,S., de
  Akinret,B., Levine,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de
  Jong,P. and Fraser,C.M.
  Mouse BAC End Sequences from Library RPCI-23
  Unpublished (1999)
  Other GSSs: RPCI-23-336P9.TU
  Contact: Shaying Zhao
  Department of Eukaryotic Genomics
  The Institute for Genomic Research
  9712 Medical Center Dr., Rockville, MD 20850, USA
  Tel: 301 838 0200
  Fax: 301 838 0208
  Email: szhao@tigr.org
  Clones are derived from the mouse BAC library RPCI-23. For BAC
  library availability, please contact Pieter de Jong
  (pieterdejong.med.buffalo.edu). Clones may be purchased from
  BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
  or from Resea ch Genetics (info@resgen.com). BAC end page:
  http://www.tigr.org/cdb/bac_end/mouse/bac_end_intro.html
  Plate: 336 row: P column: 9
  Seq primer: T7
  Class: BAC ends.
FEATURES
  source
  1..171
  /organism="Mus musculus"
  /mol_type="genomic DNA"
  /strain="C57BL/6J"
  /db_xref="taxon:10090"
  /clone="RPCI-23-336P9"
  /sex="Female"
  /lab_host="DH10B"
  /clone_lib="RPCI-23"
  /notes="Organ: Kidney/Brain; Vector: pBACe3.6; Site 1:
  EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or
  brain genomic DNA was isolated and partially digested
  with a combination of EcoRI and EcoRI Methylase. Size
  selected DNA was cloned into the pBACe3.6 vector at the
  EcoRI sites. The ligation products were transformed into
  DH10B electrocompetent cells (BRL Life Technologies)."
ORIGIN
Alignment Scores:
Pred. No.: 3.8e+03 Length: 171
Score: 27.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0
US-10-014-101B-41 (1-6) x AZ023516 (1-171)
QY 1 IleThrArgAlaArgIle 6
Db 16 ATCACCCGCTAGGATT 33
RESULT 9
CB280935
LOCUS
DEFINITION
  jaa05d01.y1 Anolis sagrei limb bud 1 Anolis sagrei cDNA 5', mRNA
  sequence.
ACCESSION
  CB280935
VERSION
  jaa05d01.y1
KEYWORDS
  GSS.
SOURCE
  Anolis sagrei (brown anole)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Lepidosauria; Squamata; Iguania; Iguanidae; Polychrotinae; Anolis.
REFERENCE
  1 (bases 1 to 179)
AUTHORS
  Losos,J., Gibson-Brown,J., Sanger,T.J., Langwith,S., Murphy,C.,
  Clifton,S., Pape,D., Marra,M., Hillier,L., Martin,J., Wylie,T.,
  Dante,M., Theising,B., Bowers,Y., Gibbons,M., Ritter,E.,
  Bennett,J., Ronko,I., Teagareishvili,R., Maguire,L., Kennedy,S.,
  Waterston,R. and Wilson,R.
  Packard/Washington University Lizard Limb Bud EST Project
  Unpublished (2003)
  Contact: Jeremy Gibson-Brown
  Packard/Washington University Lizard Limb Bud EST Project
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: est@watson.wustl.edu
  Library was constructed by Jeremy Gibson-Brown DNA sequencing by:
  Washington University Genome Sequencing Center For information on
  obtaining a clone please contact: Jeremy Gibson-Brown
  (gibbro@biology.wustl.edu)
  Putative full length read
  vector to vector length is
  Seq primer: T7.
FEATURES
  Location/Qualifiers
  1..179
  /organism="Anolis sagrei"
  /mol_type="mRNA"
  /db_xref="taxon:38937"
  /lab_host="DH10B"
  /clone_lib="Anolis sagrei limb bud 1"
  /notes="Vector: Bluescript SK+; Site_1: EcoRI; Site_2:
  BamHI; 1st strand primed with an oligo (dT) primer: double
  stranded cDNA was cloned into the EcoRI and BamHI sites of
  the pBluescript SK+ vector. Primary library;
  non-amplified."
ORIGIN
Alignment Scores:
Pred. No.: 3.97e+03 Length: 179
Score: 27.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-014-101B-41 (1-6) x CB280935 (1-179)
QY 1 IleThrArgAlaArgIle 6
Db 24 ATCACGAGAGCCAGAATA 41
RESULT 10
CC798738/c
LOCUS
DEFINITION
  CC798738 183 bp DNA linear GSS 01-JUL-2003
  SALK_147359.37.35.n Arabidopsis thaliana TDNA insertion lines
  Arabidopsis thaliana genomic clone SALK_147359.37.35.n, genomic
  survey sequence.
ACCESSION
  CC798738
VERSION
  CC798738.1 GI:32393961
KEYWORDS
  GSS.
SOURCE
  Arabidopsis thaliana (thale cress)
ORGANISM
  Arabidopsis thaliana
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE
  1 (bases 1 to 183)
AUTHORS
  Alonso,J.M., Leisner,T.J., Barajas,P., Chen,H., Cheuk,R.,

```

Gadrinab, C., Jeske, A., Karnes, M., Kim, C.J., Parker, H., Prednits, L., Shinn, P., Zimmerman, J. and Ecker, J.R.  
 A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome  
 Unpublished (2001)  
 Contact: Joseph R. Ecker  
 Salk Institute Genomic Analysis Laboratory (SIGNAL)  
 The Salk Institute for Biological Studies  
 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA  
 Tel: 858 453 4100 x1752  
 Fax: 858 558 6379  
 Email: ecker@salk.edu  
 This is single pass sequence recovered from the left border of TDNA. This sequence lies within an annotated intron of At4g21390.  
 Class: TDNA tagged.

#### FEATURES

Location/Qualifiers  
 1. 183

/organism="Arabidopsis thaliana"  
 /mol\_type="genomic DNA"  
 /ecotype="Col-0"  
 /db\_xref="taxon:3702"  
 /clone="SALK\_147359.37.35.n"

/clone\_lib="Arabidopsis thaliana TDNA insertion lines"  
 /notes="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at [http://signal.salk.edu/tdna\\_protocols.html](http://signal.salk.edu/tdna_protocols.html)"

#### ORIGIN

Alignment Scores:  
 Pred. No.: 4,06e+03 Length: 183  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-10-014-101B-41 (1-6) x CC798738 (1-183)

QY 1 IleThrArgAlaArgile 6

Db 72 ATAAACAGAGCTCGATT 55

RESULT 11  
 CA485620/c  
 LOCUS WHE4321\_A02\_A03ZS Wheat meiotic anther cDNA library Triticum  
 DEFINITION aestivum cDNA clone WHE4321\_A02\_A03, mRNA sequence.

ACCESSION CA485620  
 VERSION CA485620.1 GI:24979625  
 KEYWORDS EST.  
 SOURCE Triticum aestivum (bread wheat)

#### ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Triticum.

1 (bases 1 to 186)

Anderson, O.D., Chao, S., Crossman, C., Langridge, P., Lazo, G.R., Pham, J., Rausch, C.J., Sutton, T., Woo, J. and Wilson, C.

The structure and function of the expressed portion of the wheat genomes - Meiotic anther cDNA library  
 Unpublished (2002)

Contact: Olin Anderson

US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center  
 800 Buchanan Street, Albany, CA 94710, USA

Tel: 5105595773

Fax: 5105595818

Email: oanderson@w.usda.gov

Sequences have been trimmed to remove vector sequence and low quality sequence with phred score less than 20

#### FEATURES

source

Seq primer: SK primer.

Location/Qualifiers

1. 186

/organism="Triticum aestivum"

/mol\_type="mRNA"

/cultivar="Chinese Spring"

/db\_xref="taxon:4565"

/clone="WHE4321\_A02\_A03"

/tissue\_type="Anther"

/dev\_stage="Meiotic stages pre-meiosis-metaphase I"

/lab\_host="E. coli DH10B"

/clone\_lib="Wheat meiotic anther cDNA library"

/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; Plants were grown in a glasshouse. Anther meiotic stage was determined by removing anthers from individual primary florets. One anther was sacrificed for microscopic staging, and if determined to be between (and including) meiotic stages pre-meiosis and metaphase I, the remaining two anthers were collected and pooled for library construction. The tissue, total RNA, and poly(A) RNA were prepared, cDNA synthesised, and directionally ligated into pSPORT1 by Tim Sutton in the P Langridge Lab at the Department of Plant Science, University of Adelaide, Waite Campus, Australia. Average insert size 1.5KB. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

#### ORIGIN

Alignment Scores:  
 Pred. No.: 4,12e+03 Length: 186  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-10-014-101B-41 (1-6) x CA485620 (1-186)

QY 1 IleThrArgAlaArgile 6

Db 79 ATCACTCGCGCAGATC 62

RESULT 12

CA452616

LOCUS

DEFINITION

CA452616 194 bp mRNA linear EST 12-NOV-2002

H95-Rp1-KrIn Zea mays cDNA clone KrIn-4\_F01, mRNA sequence.

ACCESSION CA452616

VERSION CA452616.1 GI:24934398

KEYWORDS EST.

SOURCE Zea mays

#### ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD

1 (bases 1 to 194)

Bai, J., Fellers, J.P., Leach, J.E. and Hulbert, S.H.

Comparison of pathogen induced defense gene profiles on maize lines with different resistance genes  
 Unpublished (2003)

Contact: Bai J

Department of Plant Pathology

Kansas State University

4024 Throckmorton Plant Sciences Center, Manhattan, KS 66506, USA

Tel: 785-532-2328

Fax: 785-532-5692

Email: jianfa@plantpath.ksu.edu

Seq primer: T7.

#### FEATURES

source

Location/Qualifiers

1. 194

/organism="Zea mays"

/mol\_type="mRNA"

/strain="inbred line H95-Rp1-KrIn"

```

/db_xref="taxon:4577"
/clone="KrIn-4_F01"
/tissue_type="Spontaneously lesioned leaves"
/dev_stage="4 week-old plants"
/clone_lib="subtracted cDNA library of maize inbred line
H95-Rp1-KrIn"
/notes="Vector: pUC19; Suppression subtractive
hybridization; cloned into pUC19 vector. From a substracted
cDNA library of maize inbred line H95 carrying the
Rpl-KrIn gene that conditions defense response phenotype."

```

## ORIGIN

Alignment Scores:  
 Pred. No.: 4.3e+03 Length: 194  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-10-014-101B-41 (1-6) x CA452616 (1-194)

QY 1 IleThrArgAlaArgIle 6  
 |||||  
 DB 170 ATTACAGTGCACGCATC 187

## RESULT 13

CA452634  
 LOCUS KrIn-2\_B09 subtracted cDNA library of maize inbred line EST 12-NOV-2002  
 DEFINITION H95-Rp1-KrIn Zea mays cDNA clone KrIn-2\_B09, mRNA sequence.

ACCESSION CA452634  
 VERSION CA452634.1 GI:24934416  
 KEYWORDS EST.

SOURCE Zea mays  
 ORGANISM Zea mays

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.

AUTHORS Bai, J., Fellers, J.P., Leach, J.E. and Hulbert, S.H.

TITLE Comparison of pathogen induced defense gene profiles on maize lines  
 with different resistance genes

JOURNAL Unpublished (2003)

COMMENT Contact: Bai J  
 Department of Plant Pathology  
 Kansas State University  
 4024 Throckmorton Plant Sciences Center, Manhattan, KS 66506, USA  
 Tel: 785-532-2328

Fax: 785-532-5692

Email: jianfa@plantpath.ksu.edu

Seq primer: T7.

## FEATURES

source

Location/Qualifiers

```

1..195
/organism="Zea mays"
/mol_type="mRNA"
/strain="inbred line H95-Rp1-KrIn"
/db_xref="taxon:4577"
/clone="KrIn-2_B09"
/tissue_type="Spontaneously lesioned leaves"
/dev_stage="4 week-old plants"
/clone_lib="subtracted cDNA library of maize inbred line
H95-Rp1-KrIn"
/notes="Vector: pUC19; Suppression subtractive
hybridization; cloned into pUC19 vector. From a substracted
cDNA library of maize inbred line H95 carrying the
Rpl-KrIn gene that conditions defense response phenotype."

```

## ORIGIN

Alignment Scores:  
 Pred. No.: 4.32e+03 Length: 195  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-10-014-101B-41 (1-6) x CA452634 (1-195)

QY 1 IleThrArgAlaArgIle 6  
 |||||  
 DB 171 ATTACAGTGCACGCATC 188

## RESULT 14

BG412344  
 LOCUS OV2\_40\_G07.bl\_A002 Ovary 2 (OV2) Sorghum bicolor cDNA, mRNA  
 DEFINITION sequence.

ACCESSION BG412344  
 VERSION BG412344.1 GI:13317897  
 KEYWORDS EST.  
 SOURCE Sorghum bicolor (sorghum)

## ORGANISM

Sorghum bicolor  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Sorghum.

REFERENCE 1 (bases 1 to 214)

AUTHORS Cordonnier-Pratt, M.-M., Gingle, A., Marsala, C., Sudman, M. and  
 Pratt, L.H.

TITLE An EST database from Sorghum: ovaries of varying immature stages  
 JOURNAL Unpublished (2000)

COMMENT Contact: Cordonnier-Pratt MM

Laboratory for Genomics and Bioinformatics

The University of Georgia, Department of Plant Biology

Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA

Tel: 706 542 1860

Fax: 706 583 0210

Email: mmpratt@uga.edu

Seq primer: JEN REV

High quality sequence stop: 33

POLYA=No.

## FEATURES

source

Location/Qualifiers

```

1..214
/organism="Sorghum bicolor"
/mol_type="mRNA"
/db_xref="taxon:4558"
/clone_lib="Ovary 2 (OV2)"
/notes="Organ: Mix of ovaries of varying immature stages
from 8-week-old plants; Vector: pBluescript II from Lambda
Zap II; Site 1: XhoI; Site 2: EcoRI; The library was made
from poly-A RNA in the cloning vector lambda ZAP II.
Clones to be sequenced were prepared by mass excision."

```

## ORIGIN

Alignment Scores:

Pred. No.: 4.73e+03 Length: 214  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0

US-10-014-101B-41 (1-6) x BG412344 (1-214)

QY 1 IleThrArgAlaArgIle 6  
 |||||  
 DB 149 ATCACCCGTGCCAGGATT 166

## RESULT 15

BH868282  
 LOCUS hk23h11.x7 WGS-ZmaysF (JM107 adapted methyl filtered) Zea mays  
 DEFINITION genomic clone hk23h11 5', genomic survey sequence.

ACCESSION BH868282

VERSION BH868282.1 GI:22104179

KEYWORDS GSS.

**SOURCE**  
ORGANISM  
Zea mays  
Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 224)  
Katinowicz,P.D., O'Shaughnessy,A.L., Ballia,V., Dedhia,N., Ratzburger,F., King,L., Miller,B., Muller,S., Nascimento,L., Zutavern,T., McCombie,W.R. and Martienssen,R.A.  
Genomic shotgun sequences from Zea mays (methyl-filtered)  
Unpublished (2002)  
Contact: W. Richard McCombie  
Lita Auerberg Hazen Genome Sequencing Center  
Cold Spring Harbor Laboratory  
PO Box 100, Cold Spring Harbor, NY 11724, USA  
Tel: 516 367 8884  
Fax: 516 367 8874  
Email: mcombie@cshl.org  
Plate: hk23 row: h column: 11  
Seq primer: -21M13UnivRev  
Class: shotgun  
High quality sequence stop: 224.  
Location/Qualifiers  
1..224  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/cultivar="B73"  
/db\_xref="taxon:4577"  
/clone="hk23h1"  
/lab\_host="JM107 or DH5a"  
/clone\_lib="WGS-Zmav9f (JM107 adapted methyl filtered)"  
/note="Organ: immature ears; Site 1: Xba I; Site 2: Xba I; The vector was digested with XbaI and one nucleotide was added by fill in in the recessive 3' end. The genomic DNA was nebulized, end repaired, adaptor ligated and size fractionated using sephadex. The resulting fragments were between 0.8 and 3 kb and were cloned into the vector (x/y reads in M13mp19, b/g reads in pUC19). The same ligation was transformed in either JM107 or DH5a."

**FEATURES**  
source  
1..233  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="C130048M07"  
/sex="mixed"  
/tissue\_type="head"  
/dev\_stage="16 days embryo"  
/lab\_host="DH10B"  
/clone\_lib="RIKEN full-length enriched, 16 days embryo head"  
/note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGAGGATCCAGAGCTCTTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGAGATCTCGAGTTAATTAATTAATTCCTCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC 1"

**ORIGIN**  
Alignment Scores:  
Pred. No.: 233  
Score: 5,14e+03 Length: 233  
Percent Similarity: 100.00% Matches: 6  
Best Local Similarity: 100.00% Conservat: 0  
Query Match: 100.00% Mismatches: 0  
DB: 2 Indels: 0 Gaps: 0  
US-10-014-101B-41 (1-6) x BH868282 (1-224)  
Qy 1 IleThrArgAlaArgIle 6  
Db .188 ATACCCAGGCCAGGATC 205  
BB369534 233 bp mRNA linear EST 12-JUL-2000  
BB369534 RIKEN full-length enriched, 16 days embryo head Mus  
musculus cDNA clone C130048M07 3', mRNA sequence.  
BB369534  
BB369534.1 GI:9081363  
EST.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 233)  
Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Ishii,F., Ishikawa,J., Ishikawa,T., Itoh,M., Iizawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kiyosawa,H., Koike,Y., Kondo,S., Koyama,S., Kurihara,C., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tomimaga,N., Toya,T., Teunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.  
RIKEN Mouse ESTs (Konno,H., et al.)  
Unpublished (2000)  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/  
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
Thermotabilization and thermostabilization of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
Itoh,M., Katsunai,T., Akiyama,J., Shibata,K., Iizawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.  
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)  
Carninci,P. and Hayashizaki,Y.  
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)  
Please visit our web site (http://genome.rtc.riken.go.jp) for further details.  
Location/Qualifiers  
1..233  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="C130048M07"  
/sex="mixed"  
/tissue\_type="head"  
/dev\_stage="16 days embryo"  
/lab\_host="DH10B"  
/clone\_lib="RIKEN full-length enriched, 16 days embryo head"  
/note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGGATCCAGAGCTCTTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGATCTCGAGTTAATTAATTAATTCCTCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC 1"

**REFERENCE**  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

```

QY      1 IleThrArgAlaArgile 6
Db      34 ATTACAGAGCCCGTATC 51

RESULT 17
BE764604/c
LOCUS   BE764604
DEFINITION 235 bp mRNA linear EST 19-SEP-2000
ACCESSION FM2-NT0075-150500-002-b01 NT0075 Homo sapiens cDNA, mRNA sequence.
VERSION   BE764604
KEYWORDS  BE764604.1 GI:10194528
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
          Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
          Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
          Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
          O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
          Simpson,A.J.
          Shotgun sequencing of the human transcriptome with ORF expressed
          sequence tags
          Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
          20202663
          10737800
          Contact: Simpson A.J.G.
          Laboratory of Cancer Genetics
          Ludwig Institute for Cancer Research
          Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
          Brazil
          Tel: +55-11-2704922
          Fax: +55-11-2707001
          Email: asimpson@ludwig.org.br
          This sequence was derived from the FAPESP/LICR Human Cancer Genome
          Project. This entry can be seen in the following URL
          (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=PM2-NT0075-150
          500-002-b01&t3=2000-05-15&t4=1)
          Seq primer: puc 18 forward
          High quality sequence start: 33
          High quality sequence stop: 151.
FEATURES
        source
        1..235
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /dev_stage="Adult"
        /clone_lib="NT0075"
        /notes="Organ: nervous tumor; Vector: puc18; Site_1: SmaI;
        Site_2: SmaI; A mini-library was made by cloning products
        derived from ORESTES PCR (U.S. Letters Patent application
        No. 196,716 - Ludwig Institute for Cancer Research)
        profiles into the puc 18 vector. Reverse transcription of
        tissue mRNA and cDNA amplification were performed under
        low stringency conditions."
ORIGIN
Alignment Scores:
Pred. No.: 5.19e+03 Length: 235
Score: 27.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-014-101B-41 (1-6) x BE764604 (1-235)

QY      1 IleThrArgAlaArgile 6
Db      108 ATCACCAGGCGCCGGATC 91

```

```

RESULT 18
BI190218
LOCUS   BI190218
DEFINITION 237 bp mRNA linear EST 10-JUL-2001
          hlf02fs.f1 Fusarium sporotrichioides Tri 10 overexpressed cDNA
          library Fusarium sporotrichioides cDNA clone hlf02fs 3', mRNA
          sequence.
ACCESSION BI190218
VERSION   BI190218.1 GI:14663897
KEYWORDS  EST.
SOURCE    Fusarium sporotrichioides
ORGANISM  Fusarium sporotrichioides
REFERENCE Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
AUTHORS   Hypocreomycetidae; Hypocreales; mitosporic Hypocreales; Fusarium.
          1 (bases 1 to 237)
          Ren,Q., Tag,A., Feplow,A., Lai,H., Kupfer,C., Peterson,A.,
          Beremand,M. and Roe,B.
          Analysis of a Fusarium sporotrichioides EST database
          Unpublished (2001)
          Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
          Department of Chemistry and Biochemistry
          Advanced Center for Genome Technology, University of Oklahoma
          620 Parrington Oval, Norman, OK 73019, USA
          Tel: 405 325 4912
          Fax: 405 325 7762
          Email: broe@ou.edu
          Contact Dr. Marian Beremand regarding clone availability Included
          is the best homolog from a blastx search of Genbank nr 04-09-01
          71 1.3 gi|323156|pir||S3215 protein-tyrosine kinase (EC 2.7.1.112)
          GCT
          Seq primer: M13-20
          High quality sequence stop: 79.
FEATURES
        source
        1..237
        /organism="Fusarium sporotrichioides"
        /mol_type="mRNA"
        /strain="Tri 10"
        /db_xref="taxon:5514"
        /clone_lib="hlf02fs"
        /clone="hlf02fs"
        /notes="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
        XhoI; 5' end of cDNA cloned into EcoRI site of
        pBluescript; 3' end of cDNA cloned into XhoI site of
        pBluescript"
ORIGIN
Alignment Scores:
Pred. No.: 5.23e+03 Length: 237
Score: 27.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-014-101B-41 (1-6) x BI190218 (1-237)

QY      1 IleThrArgAlaArgile 6
Db      191 ATAACAAGAGCTCGAATC 208

RESULT 19
BG994575/c
LOCUS   BG994575
DEFINITION 239 bp mRNA linear EST 13-JUN-2001
ACCESSION RC4-HT1141-130201-011-d06 HT1141 Homo sapiens cDNA, mRNA sequence.
VERSION   BG994575
KEYWORDS  BG994575.1 GI:14398645
SOURCE    EST.
ORGANISM  Homo sapiens (human)
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

```

Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J., and Simpson, A.J.G.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20020663

10737800

# COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC4&t2=RC4-HT1141-130201-011-d06&t3=2001-02-13&t4=1>)

Seq primer: puc 18 forward

High quality sequence start: 108

High quality sequence stop: 159.

## FEATURES

source

1..239

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

/clone\_lib="HT1141"

/note="Organ: head neck; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

## ORIGIN

Alignment Scores:  
Pred. No.: 5.27e+03 Length: 239  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-41 (1-6) x BG994575 (1-239)

QY 1 IleThrArgAlaArgile 6

|||||

Db 207 ATTACCCGGCCGGTATA 190

## RESULT 20

BE025443/c

LOCUS

DEFINITION 945028H10.Y2 945 - Mixed adult tissues from Walbot lab, same as 707

(SK) Zea mays cDNA, mRNA sequence.

ACCESSION BE025443

VERSION BE025443.1

KEYWORDS EST.

SOURCE Zea mays

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 249)

REFERENCE Walbot, V.

AUTHORS Zea mays

TITLE Maize ESTs from various cDNA libraries sequenced at Stanford

UNPUBLISHED (1999)

## COMMENT

Contact: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Plate: 945028 row: H column: 10.

## FEATURES

source

1..249

/organism="Zea mays"

/mol\_type="mRNA"

/cultivar="W23"

/db\_xref="taxon:4577"

/tissue\_type="tassel, kernal, silk, husk, root, leaf"

/dev\_stage="fully-grown"

/lab\_host="DH10B"

/clone\_lib="945 - Mixed adult tissues from Walbot lab,

same as 707 (SK)"

/note="Organ: tassel, kernal, silk, husk, root, leaf;

Vector: pGAD10; Site 1: EcoRI; cDNA library from fully

differentiated maize tissues from an active Mutator plant.

Tissue ratio is 4:2:1:1:1 (tassel, kernel, silk, husk,

root, leaf). Unidirectionally cloned. New library number

given to library 707 for additional sequencing."

## ORIGIN

Alignment Scores:  
Pred. No.: 5.49e+03 Length: 249  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0

US-10-014-101B-41 (1-6) x BE025443 (1-249)

QY 1 IleThrArgAlaArgile 6

|||||

Db 50 ATTACAGTGCACGCATT 33

## RESULT 21

CD712578

LOCUS

DEFINITION CD712578 253 bp mRNA linear EST 25-JUN-2003

VVB021G12 403503 An expressed sequence tag database for abiotic

stressed leaves of Vitis vinifera var. Chardonnay Vitis vinifera

cDNA clone VVB021G12 5, mRNA sequence.

ACCESSION CD712578

VERSION CD712578.1

KEYWORDS EST.

SOURCE Vitis vinifera

ORGANISM Vitis vinifera

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; Vitaceae; Vitis.

1 (bases 1 to 253)

REFERENCE Cramer, G.R. and Cushman, J.C.

AUTORS An expressed sequence tag database for abiotic stressed leaves of

Vitis vinifera var. Chardonnay

UNPUBLISHED (2002)

CONTACT: Cushman JC

DEPARTMENT OF Biochemistry

UNIVERSITY OF Nevada

MS200, Reno, NV 89557-0014, USA

TEL: 775-784-1918

FAX: 775-784-1650

EMAIL: jcushman@unr.edu

PCR PRIMERS

FORWARD: T3 20mer

BACKWARD: T7 21mer (backward)

PLATE: 0 row: G column: 12

Seq primer: T7 20mer (forward)

High quality sequence stop: 253.

FEATURES  
source

Location/Qualifiers  
1. .253  
/organism="Vitis vinifera"  
/mol\_type="mRNA"  
/db\_xref="taxon:29760"  
/clone="VVB021G12"  
/tissue\_type="leaf"  
/dev\_stage="juvenile and adult"  
/clone\_lib="An expressed sequence tag database for abiotic stressed leaves of Vitis vinifera var. Chardonnay"  
/notes="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site 1: EcoRI; Site 2: XhoI; Library construction was performed according to Stratagene's recommended protocol for the Lambda UnizapR vector and cDNA synthesis kit."

## ORIGIN

Alignment Scores:  
Pred. No.: 5.58e+03 Length: 253  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-014-101B-41 (1-6) x CD712578 (1-253)

QY 1 IleThrArgAlaArgile 6  
|||||  
DB 143 ATTACAGAGCAGGATA 160

## RESULT 22

BM078603/c 258 bp mRNA linear EST 14-NOV-2001  
LOCUS MEST122-A01.T3 ISUM4-TN Zea mays cDNA clone MEST122-A01 3', mRNA  
DEFINITION sequence.

ACCESSION BM078603  
VERSION BM078603.1 GI:16925535  
KEYWORDS EST.  
SOURCE Zea mays

## ORGANISM

Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 258)  
Qiu, F., Cui, F., Guo, L., Ashlock, D.A., Wen, T.J. and Schnable, P.S.  
Expressed Sequence Tags from B73 Maize Seedlings and Silks  
Unpublished (2001)

## JOURNAL

COMMENT  
Contact: Patrick S. Schnable  
Schnable Laboratory  
Iowa State University  
G405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA  
Tel: 515-294-0975  
Fax: 515-294-2299  
Email: schnable@iastate.edu

Phred software,  
(<http://depts.washington.edu/ventures/collabtr/direct/index.htm#b>  
rt). Overall sequence quality assessment and vector trimming were conducted using the Lucy software (<http://www.tigr.org/softlab/lucy/>). Lucy parameters were set to ensure an overall trimmed quality of 97.5% or better without any vector fragments in the chosen high-quality region of each sequence. Low-quality bases between the poly-T and the high-quality region were replaced with N's to serve as spacers.

PCR Primers  
FORWARD: Forward PCR primer sequence, primer T7-1 (AA TAC GAC TCA CTA TAG)  
BACKWARD: Backward PCR primer sequence, primer T3 (ATT AAC CCT CAC TAA AG)

Seq primer: primer T3 (ATT AAC CCT CAC TAA AG).

FEATURES  
source

Location/Qualifiers  
1. .258  
/organism="Zea mays"

/mol\_type="mRNA"  
/cultivar="B73"  
/db\_xref="taxon:4577"  
/clone="MEST122-A01"  
/tissue\_type="Seedling and silk"  
/lab\_host="DH10B"  
/clone\_lib="ISUM4-TN"

/notes="Vector: pT73PAC; Site 1: EcoRI; Site 2: NotI; ds-cDNA molecules were generated as follows. First-strand cDNA was prepared from oligo-dT selected mRNA by priming with a NotI oligo-dT primer (5' AACTGGAAGAATTCGGCGCGCAGGAATTTTTTTTTTTT). The resulting DNA:RNA hybrid was treated with RNase H and used as a template for DNA PolI-catalyzed second strand synthesis. After the addition of EcoRI adaptors, the ds-cDNAs were digested with NotI and size-selected. The resulting molecules were directionally cloned into the EcoRI and NotI sites of the pT73PAC vector. The library then went through one round of normalization to Cot value of 5 based on the methods of Marcelino Bento Soares (Genome Research 6: 791-806, 1996)."

## ORIGIN

Alignment Scores:  
Pred. No.: 5.69e+03 Length: 258  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-41 (1-6) x BM078603 (1-258)

QY 1 IleThrArgAlaArgile 6  
|||||  
DB 143 ATTACAGTCGACGCATT 126

## RESULT 23

AV323709 259 bp mRNA linear EST 09-NOV-1999  
AV323709 RIKEN full-length enriched, 11 days embryo head Mus musculus cDNA clone 6230412A13 3' similar to D10916 Mouse mRNA of NEDD-6 gene, 3'terminal sequence, mRNA, mRNA sequence.

ACCESSION AV323709.1 GI:6293626  
VERSION  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)

## ORGANISM

Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 259)

## REFERENCE

AUTHORS  
Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F., Ishii, Y., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kigawa, I., Kai, C., Kawai, J., Kikuchi, N., Kojima, Y., Koya, S., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K., Shibata, Y., Shigenoto, Y., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Takahashi, F., Tateno, M., Tominaga, N., Tsunoda, Y., Watanabe, S., Yamamura, T., Yasunishi, A., Yokota, T., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.  
RIKEN Mouse ESTs (Konno, H., et al. 1999)  
Unpublished (1999)

## TITLE

Unpublished (1999)

## JOURNAL

COMMENT  
Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216

Email: genome-resescsc.riken.jp, URL: <http://genome.gsc.riken.jp/>  
Sasaki, N., Izawa, M., Wataniki, M., Ozawa, K., Tanaka, T., Yoneda, Y., Matsuura, S., Carninci, P., Muramatsu, M., Okazaki, Y. and

Hayashizaki, Y.  
Transcriptional sequencing: A method for DNA sequencing using RNA polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)  
Itoh, M., Kitanai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y., and Hayashizaki, Y.  
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)  
Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES  
source  
1. .259  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="6230412A13"  
/sex="mixed"  
/tissue type="head"  
/dev stage="11 days embryo"  
/lab\_host="DH108"  
/clone lib="RIKEN full-length enriched, 11 days embryo head"  
/notes="Site 1: Sali; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'  
GAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCGAGTAAATTAATCCGCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda PLC I. Cloning sites, 5' end: Sali; 3' end: BamHI."

## ORIGIN

Alignment Scores:  
Pred. No.: 5.71e+03 Length: 259  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 1 Gaps: 0

US-10-014-101B-41 (1-6) x AV323709 (1-259)

QY 1 IleThrArgAlaArgile 6

Db 14 ATCACAGGGCCCGAATC 31

## RESULT 24

AW620066  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS

AW620066  
819 MARC PBE Sus scrofa cDNA 5', mRNA sequence.  
EST.  
GI:7326250  
Sus scrofa (pig)  
Sus scrofa  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
1 (bases 1 to 260)  
Smith, T.P.L., Fahrenkrug, S.C., Rohrer, G.A., Simmen, F.A.,

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT

Rexroad, C.E. and Keele, J.W.  
Mapping of expressed sequence tags from a porcine early embryonic cDNA library  
Anim. Genet. 32 (2), 66-72 (2001)  
21314990  
11421940  
Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: [smith@email.marc.usda.gov](mailto:smith@email.marc.usda.gov)  
Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross\_match with the -minscore 18 and -minmatch 12 options.

PCR Primers  
FORWARD: GGAAACAGCTATGACCATG  
BACKWARD: GTAAACGACGGCCAGT  
Seq primer: AATTACCTCTACTAAAGG.

FEATURES  
source

1. .260  
Location/Qualifiers  
/organism="Sus scrofa"  
/mol\_type="mRNA"  
/db\_xref="taxon:9823"  
/tissue type="Day 12 whole embryos"  
/lab\_host="XLOLR"  
/clone lib="MARC PBE"  
/notes="Vector: pBLUESCRIPT SK-; Site 1: EcoRI; Site 2: XhoI; Library made from pool of embryos in spherical and filamentous stages of development (7.5% and 92.5%, respectively, of each stage) as described in Choi et al, Endocrinology 137, 1457-67, 1996."

## ORIGIN

Alignment Scores:  
Pred. No.: 5.73e+03 Length: 260  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0

US-10-014-101B-41 (1-6) x AW620066 (1-260)

QY 1 IleThrArgAlaArgile 6

Db 153 ATTACAGGGCCCGAATC 170

## RESULT 25

BI055989/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS

BI055989  
PM2-GN0431-150201-001-c03 GN0431 Homo sapiens cDNA, mRNA sequence.  
EST.  
GI:14463519  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 261)  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
10737800  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics

## TITLE

JOURNAL  
MEDLINE  
PUBMED  
COMMENT

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
10737800  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM2&t2=PM2-GN0431-  
150201-001-c03&t3=2001-02-15&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 26  
High quality sequence stop: 261.

## FEATURES

source

1. 261  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="GN0431"  
/notes="Organ: placenta normal; Vector: puc18; Site 1:  
SmaI; Site 2: SmaI; A mini-library was made by cloning  
products derived from ORESTES PCR (U.S. Letters Patent  
application No. 196, 716 - Ludwig Institute for Cancer  
Research) profiles into the pUC 18 vector. Reverse  
transcription of tissue mRNA and cDNA amplification were  
performed under low stringency conditions."

## ORIGIN

Alignment Scores:  
Pred. No.: 5.75e+03 Length: 261  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-10-014-101B-41 (1-6) x BI055989 (1-261)

QY 1 IleThrArgAlaArgIle 6

Db 228 ATAAACCGGCGCGGTATA 211

## RESULT 26

CA811112

LOCUS

DEFINITION CA22LI041R-D10 Cabernet Sauvignon Leaf - CA22LI Vitis vinifera cDNA

ACCESSION

CA811112

VERSION

CA811112.1

KEYWORDS

EST.

SOURCE

Vitis vinifera

ORGANISM

Vitis vinifera

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CAES Genome Facility

UC Davis Department of Plant Pathology

1 Shields Ave., Davis, CA 95616, USA

Tel: 530 754 6561

Fax: 530 754 6617

Email: drcook@ucdavis.edu

Seq primer: GCCAACGAATGGTCTAG.

Location/Qualifiers

1. 261

/organism="Vitis vinifera"

/mol\_type="mRNA"  
/cultivar="Cabernet Sauvignon"  
/db\_xref="taxon:29760"  
/clone="CA22LI041R-D10"  
/sex="hermaphrodite"  
/dev\_stage="Late season sample"  
/lab\_host="DH5alpha"  
/clone\_lib="Cabernet Sauvignon Leaf - CA22LI"  
/notes="Organ: Leaf; Vector: pDNR; Site 1: SfiI; Site 2:  
SfiI; CA22LI is a cDNA library of Cabernet Sauvignon  
leaves. The leaves were collected on September 20, 2001,  
in Napa Valley, California, and represent leaves in late  
season development. These leaves were verified to be  
infected with the bacterial pathogen, Xylella fastidiosa,  
based on a diagnostic assay using PCR and Xylella-specific  
primer pairs. Plants were symptomatic at the time of  
collection. cDNA made by oligo-dT priming and  
directionally cloned. 5' and 3' adaptors were used in  
cloning as follows: 5'-AAGCAGTGTATCAACGACAGTGGCCATTACGGCGG-3' and  
5'-ATTCTAGCGCGGCGCGACATG-dT(30)NN-3'. Library was  
constructed using the Clontech Creator SMART kit and  
size-selected to contain the 0.5-3 kb size fraction."

## ORIGIN

Alignment Scores:  
Pred. No.: 5.75e+03 Length: 261  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-014-101B-41 (1-6) x CA811112 (1-261)

QY 1 IleThrArgAlaArgIle 6

Db 201 ATTACAGAGACCGGATA 218

## RESULT 27

BB319596

LOCUS

DEFINITION

BB319596

ACCESSION

BB319596

VERSION

BB319596.1

KEYWORDS

EST.

SOURCE

Mus musculus

ORGANISM

Mus musculus

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (2000)

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

BB319596 262 bp mRNA linear EST 11-JUL-2000  
BB319596 RIKEN full-length cDNA clone B230381C22 3', mRNA sequence.  
Quadrigenina Mus musculus cDNA clone B230381C22 3', mRNA sequence.

BB319596

BB319596.1

GI:9026631

Mus musculus (house mouse)

Mus musculus

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 262)

Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T.,

Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,

Hirose, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,

Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,

Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C.,

Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H.,

Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K.,

Shibata, K., Shibata, Y., Shigemoto, Y., Shinagawa, A., Shiraki, T.,

Sogabe, Y., Sugahara, Y., Suzuki, H., Teunoda, Y., Teyoda, Y., Teyoda, Y.,

Takahashi, F., Tominaga, N., Toyama, T., Yamanaka, I., Yano, R., Yasunishi, A.,

Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and

Hayashizaki, Y.

RIKEN Mouse ESTs (Konno, H., et al.)

Unpublished (2000)

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216  
Email: genome.resgsc.riken.jp, URL: http://genome.gsc.riken.jp/  
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S.,  
Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Thermostabilization and thermoactivation of thermolabile enzymes by  
trehalose and its application for the synthesis of full length  
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
Tomaru, Y., Carninci, P., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,  
Okazaki, Y. and Hayashizaki, Y.  
Automated filtration-based high-throughput plasmid preparation  
system. Genome Res. 9 (5), 463-470 (1999)  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,  
19-44 (1999)  
Please visit our web site (http://genome.rtc.riken.go.jp) for  
further details.

FEATURES source  
1. .262  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="B230381C22"  
/sex="male"  
/tissue\_type="corpora quadrigemina"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="RIKEN full-length enriched, adult male corpora  
quadrigemina"  
/note="Site 1: SalI; Site 2: BamHI; cDNA library was  
prepared and sequenced in Mouse Genome Encyclopedia  
Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in  
RIKEN. Division of Experimental Animal Research in Riken  
contributed to prepare mouse tissues. 1st strand cDNA was  
primed with a primer [5',  
GAGAGAGAGAGATCCAGAGCTCTTTTTTTTTTTTNN 3'], cDNA was  
prepared by using trehalose thermo-activated reverse  
transcriptase and subsequently enriched for full-length by  
cap-trapper. cDNA went through one round of normalization  
to Rot = 20.0 and subtraction to Rot = 459.0. Second  
strand cDNA was prepared with the primer adapter of  
sequence [5' GAGAGAGAGATTCGAGTTAATTAAATTAATCCCCCCCCC  
3']. cDNA was cleaved with XhoI and BamHI. Vector: a  
modified pBluescript KS(+) after bulk excision from Lambda  
FLC I."

ORIGIN  
Alignment Scores:  
Pred. No.: 5.77e+03 Length: 262  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0  
US-10-014-101B-41 (1-6) x BB319596 (1-262)  
QY 1 IleThrArgAlaArgIle 6  
DB 17 ATCACTCGGCAGAATT 34  
RESULT 28  
B81555 263 bp DNA linear GSS 25-JUN-1998  
LOCUS CIT-HSP-2060M18.TRB CIT-HSP Homo sapiens genomic clone 2060M18,  
DEFINITION genomic survey sequence.  
ACCESSION B81555  
VERSION B81555.1 GI:2868578  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 263)  
Adams, M.D., Rounsley, S.D., Zhao, S., Field, C.E., Bass, S., Linher, K.,  
Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., Shizuya, H.,  
Simon, M. and Venter, J.C.  
Use of a random BAC End Sequence Database for Sequence-Ready Map  
Building (1998)  
Unpublished (1998)  
Other GSSs: CIT-HSP-2060M18.TF  
Contact: Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: mdamas@tigr.org  
Clones are available from Research Genetics (info@resgen.com). BAC  
end search page:  
http://www.tigr.org/tdb/humgen/bac\_end\_search/bac\_end\_search.html  
Seq primer: M13 Reverse  
Class: BAC ends.  
Location/Qualifiers  
1. .263  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="GDB:7060533"  
/db\_xref="taxon:9606"  
/clone="2060M18"  
/sex="Male"  
/cell\_type="Sperm"  
/clone\_lib="CIT-HSP"  
/note="Vector: pBelobAC11; Site\_1: HindIII; Site\_2:  
HindIII"

FEATURES source  
1. .263  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="GDB:7060533"  
/db\_xref="taxon:9606"  
/clone="2060M18"  
/sex="Male"  
/cell\_type="Sperm"  
/clone\_lib="CIT-HSP"  
/note="Vector: pBelobAC11; Site\_1: HindIII; Site\_2:  
HindIII"

ORIGIN  
Alignment Scores:  
Pred. No.: 5.79e+03 Length: 263  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0  
US-10-014-101B-41 (1-6) x B81555 (1-263)  
QY 1 IleThrArgAlaArgIle 6  
DB 104 ATCAAGGGCAGAATT 121

RESULT 29  
CR516082/c  
LOCUS CR516082 Mouse pBluescript Lion Mus musculus cDNA clone  
DEFINITION LI0NP462C06364 3', mRNA sequence.  
ACCESSION CR516082  
VERSION CR516082.1 GI:49908014  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 265)  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Contact: Inge Arlart  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
Heubnerweg 6, D-14059 Berlin, Germany  
Email: www.rzpd.de  
RZPD; LI0NP462C06364.  
RZPDLIB;

Mouse ArrayTAG cDNA  
 http://www.rzpd.de/cgi-bin/products/showlib.pl.cgi?response?libNo=462 Contact: Inge Airlart  
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
 Heubnerweg 6, D-14059 Berlin, Germany  
 Tel: +49 30 32639 100  
 Fax: +49 30 32639 111  
 www.rzpd.de  
 This clone is available royalty-free from RZPD;  
 contact RZPD (clone@rzpd.de) for further information. Seq primer:  
 RP: CAGGAACACGCTATGAC.

## FEATURES

source

Location/Qualifiers  
 1..265  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="LIONP462C06364"  
 /lab\_host="DH10B"  
 /clone\_lib="Mouse pBluescript Lion"

## ORIGIN

Alignment Scores:  
 Pred. No.: 5.84e+03 Length: 265  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 7 Gaps: 0

US-10-014-101B-41 (1-6) x CR516082 (1-265)

QY 1 IleThrArgAlaArgile 6  
 |||||  
 DB 159 ATAACAGGCGCGCATC 142

## RESULT 30

BE529872  
 LOCUS M75N155TM Arabidopsis developing seed Arabidopsis thaliana cDNA  
 DEFINITION clone 600039263R1 5', mRNA sequence.

ACCESSION BE529872  
 VERSION BE529872.1 GI:9787850

KEYWORDS EST.  
 SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE 1 (bases 1 to 266)  
 AUTHORS White,J.A., Todd,J., Newman,T., Focks,N., Girke,T., Martinez de  
 Ilarduya,O., Jaworski,J.G., Ohlrogge,J. and Benning,C.

TITLE A new set of Arabidopsis expressed sequence tags from developing  
 seeds. The metabolic pathway from carbohydrates to seed oil

JOURNAL Plant Physiol. 124 (4), 1582-1594 (2000)

MEDLINE 20567808

PUBMED 11115876

COMMENT Contact: Benning, C

Dept. of Biochemistry & Molecular Biology

Michigan State University

224 Biochemistry, Michigan State University, East Lansing, MI

48824, USA

Tel: 517 355 1609

Fax: 517 353 9334

Email: benning@msu.edu

Clones were originally prepared at Michigan State University.

Arabidopsis Biological Resource Center, The Ohio State University,

309 Botany & Zoology Bldg., 1735 Neil Avenue, Columbus, OH 43210

USA, FAX: 6142920603 TEL: 6142929371.

## FEATURES

source

Location/Qualifiers  
 1..266  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /ecotype="Columbia"

/db\_xref="taxon:3702"  
 /clone="600039263R1"  
 /issue\_type="seed"  
 /dev\_stage="5-13 days after flowering"  
 /lab\_host="E.coli"  
 /clone\_lib="Arabidopsis developing seed"  
 /notes="Organ: Developing seed; Vector: pBluescript SK-;  
 Site\_1: EcoRI; Site\_2: XhoII"

## ORIGIN

Alignment Scores:  
 Pred. No.: 5.86e+03 Length: 265  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0

US-10-014-101B-41 (1-6) x BE529872 (1-266)

QY 1 IleThrArgAlaArgile 6  
 |||||  
 DB 226 ATCACTCGCGCAGCAATC 243

## RESULT 31

CK691445/c

LOCUS CK691445

DEFINITION

IMAGE:7155314 5', mRNA sequence.

ACCESSION CK691445

VERSION CK691445.1 GI:42443781

KEYWORDS EST.

SOURCE Danio rerio (zebrafish)

ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Danio.

REFERENCE 1 (bases 1 to 266)

AUTHORS Wei,C., Mathavan,S., Thoreau,H., Lim,L., Lee,C. and Ruan,Y.

TITLE Genome Institute of Singapore, Zebrafish Gene Collection

JOURNAL Unpublished (2004)

COMMENT Contact: Ruan Y

Cloning and Sequencing

Genome Institute of Singapore

60 Biopolis Street, #02-01, Genome, Singapore 138672

Tel: +65 6478 8073

Fax: +65 6478 9059

Email: ruan@igis.a-star.edu.sg

GIS Clone ID: ZF101-P00056-BR2\_G24

PCR Primers

FORWARD: M13

BACKWARD: M13

Plate: ZF101-P00056-BR2 row: G column: 24

Seq primer: CCGCATACTTGTATAGCA

High quality sequence stop: 266.

Location/Qualifiers

1..266

/organism="Danio rerio"

/mol\_type="mRNA"

/strain="Singapore local strain"

/db\_xref="taxon:7955"

/clone="IMAGE:7155314"

/issue\_type="Embryo"

/dev\_stage="7 Different embryonic stages (From just

fertilized Embryos to 72 hours just hatched baby fish)"

/lab\_host="DH10B"

/clone\_lib="GISZF001 ra"

/notes="Vector: pDNR-LIB; Site 1: Sfi A (GGCATTACGGCC);

Site 2: Sfi B (GGCCGAGCGGCC); Priming method: Sfi-(dr)30

Primed; Priming sequence:

5.ATTCTAGAGCGCGCGGCACATG(T)30VN; Directionally

cloned, 5' cloning site: Sfi A site GGCATTACGGCC; 5'

linker/adaptor sequence: 5.AAGCAGTGTATCAGCAGATGGCC;

3' cloning site: Sfi B site GCCGAGCGGCC ; 3' linker/adaptor sequence: same as the priming sequence ; Average insert size: 2kb ; For PCR insert analysis: Use M13 Forward and reverse primers ; Library Amplified ; Recombinants (inserts): 98% ; Library complexity: 5x10<sup>6</sup> ; Full-length construction (method): SMART, a Clontech method The pooled tissue RNA was collected and used to construct full length enriched cDNA library and also served as template to synthesize complex first strand cDNA probe. Two high density colony arrays were made from over 110K cDNA clones and hybridized with the probes. Low intensity clones were selected as they represented rare expressed clones. The hybridization intensities for all clones span from 0 to 1.8 million counts and the low abundant class ranged from 0 to 13,000."

## ORIGIN

Alignment Scores:  
 Pred. No.: 5.86e+03 Length: 266  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 7 Gaps: 0

US-10-014-101B-41 (1-6) x CK691445 (1-266)

QY 1 IleThrArgAlaArgile 6  
 DB 89 ATTACACGAGCGAGATC 72

## RESULT 32

BB577396

LOCUS

DEFINITION BB577396 RIKEN full-length enriched, 11 days embryo head Mus musculus cDNA clone 6230401E23 5', mRNA sequence.

ACCESSION BB577396

VERSION BB577396.1

KEYWORDS GI:11473940

SOURCE EST.

ORGANISM Mus musculus (house mouse)

REFERENCE

AUTHORS

1 (bases 1 to 267)

Aizawa,K., Akahira,S., Akimura,T., Arai,A., Arakawa,T.,

Carninci,P., Hanagaki,T., Hayatsu,N., Hiraoka,T., Hirozane,T.,

Hodoyama,Y., Inotani,K., Ishii,Y., Itoh,M., Izawa,M., Kawai,J.,

Kojima,Y., Konno,H., Kusakabe,M., Matsuyama,T., Miyazaki,A.,

Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Okazaki,Y.,

Okido,T., Owa,C., Sakai,C., Sakai,K., Sasaki,D., Sato,K.,

Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,

Suzuki,H., Tagawa,A., Takahashi,F., Tanaka,T., Toya,T.,

Watabiki,A., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshiki,A.,

Muramatsu,M. and Hayashizaki,Y.

RIKEN Mouse ESTs (Aizawa,K. et al. 2000)

Unpublished (2000)

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center(GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suenro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome.res@sc.riken.jp, URL: http://genome.gsc.riken.jp/

Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S.,

Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

Thermosensitization and thermoinactivation of thermolabile enzymes by

trehalose and its application for the synthesis of full length

cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)

Itoh,M., Kitsuai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,

Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M.,

Okazaki,Y. and Hayashizaki,Y.

Automated filtration-based high-throughput plasmid preparation

system. Genome Res. 9 (5), 463-470 (1999)  
 Carninci,P. and Hayashizaki,Y.  
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)

Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

## FEATURES

## source

Location/Qualifiers  
 1..267  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="6230401E23"  
 /sex="mixed"  
 /tissue\_type="head"  
 /dev\_stage="11 days embryo"  
 /lab\_host="DH10B"  
 /clone\_lib="RIKEN full-length enriched, 11 days embryo head"

/note="Site 1: Sall; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGATCTCGAGTAAATTAATCCGCCGCCGCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: Sall; 3' end: BamHI."

## ORIGIN

Alignment Scores:  
 Pred. No.: 5.86e+03 Length: 267  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0

US-10-014-101B-41 (1-6) x BB577396 (1-267)

QY 1 IleThrArgAlaArgile 6

DB 144 ATTACACGAGCGAGATC 161

## TITLE

JOURNAL

COMMENT

LOCUS BB577396/267 bp DNA linear GSS 05-FEB-2003  
 DEFINITION PUBK67YTD ZM 0.6 1.0 KB Zea mays genomic clone ZMMBTA078K14,

ACCESSION BB577396

VERSION BB577396.1

KEYWORDS GI:28220342

SOURCE GSS.

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 267)

Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,

Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and

Bennetzen,J.

Maize Genomics Consortium

Unpublished (2003)

Contact: Cathy Whitelaw

TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TF  
Class: sheared ends,  
Location/Qualifiers

# FEATURES

source  
1..267  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZM8P7a078K14"  
/note="lib=ZM 0.6-1.0 kb"  
/notes="vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high  
Cot selected genomic DNA library"

## ORIGIN

Alignment Scores:  
Pred. No.: 5.88e+03 Length: 267  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0

US-10-014-101B-41 (1-6) x BZ671989 (1-267)

QY 1 IleThrArgAlaArgile 6  
|||||  
DB 254 ATCACAGGCGACGATC 237

## RESULT 34

BH228153  
LOCUS  
DEFINITION BH228153 269 bp DNA linear GSS 08-NOV-2001  
1006144E12.xl 1006 - RescueMu Grid G Zea mays genomic, genomic  
survey sequence.

ACCESSION BH228153  
VERSION BH228153.1 GI:16828887

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 269)

Walbot, V.

Maize genomic sequences found using engineered RescueMu transposon

Unpublished (2001)

Contact: Walbot V

Department of Biological Sciences

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221

Email: walbot@stanford.edu

Plate: 1006144 row: 5

Class: transposon-tagged.

Location/Qualifiers

## FEATURES

source  
1..269  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/cultivar="mixed background W23/A188/B73"  
/db\_xref="taxon:4577"  
/tissue\_type="leaf"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="1006 - RescueMu Grid G"  
/note="Organ: leaf; Vector: RescueMu (engineered from  
pBluescript backbone); Site 1: BamHI; Site 2: BglII;  
RescueMu is a 4.9 kb, modified maize Mu transposon  
designed to allow plasmid rescue from total genomic DNA.

Mu elements insert preferentially into transcription  
units. For more information on RescueMu, go to the web  
site 'www.zmdb.iastate.edu' and follow the links for  
'RescueMu.' Grid G was grown at Stanford in 2000. DNA was  
extracted from leaf punches, double digested using BamHI  
and BglII, and ligated to form circular plasmids. DH10B  
cells were transformed and then screened on LB plates with  
ampicillin."

## ORIGIN

Alignment Scores:  
Pred. No.: 5.92e+03 Length: 269  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0

US-10-014-101B-41 (1-6) x BH228153 (1-269)

QY 1 IleThrArgAlaArgile 6  
|||||  
DB 161 ATAACTCGCGCCGAATC 178

## RESULT 35

CB921222/c

LOCUS

DEFINITION

CB921222 279 bp mRNA linear EST 25-APR-2003  
VVD068E05 353077 An expressed sequence tag database for abiotic  
stressed berries of Vitis vinifera var. Chardonnay Vitis vinifera  
cDNA clone VVD068E05 5, mRNA sequence.

ACCESSION CB921222

VERSION CB921222.1 GI:30135884

KEYWORDS EST.

SOURCE Vitis vinifera

ORGANISM Vitis vinifera

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; Vitaceae; Vitis.

1 (bases 1 to 279)

Cushman, J.C.

An expressed sequence tag database for abiotic stressed berries of

Vitis vinifera var. Chardonnay

Unpublished (2002)

Contact: Cushman JC

Department of Biochemistry

University of Nevada

MS200, Reno, NV 89557-0014, USA

Tel: 775-784-1918

Fax: 775-784-1650

Email: jcushman@unr.edu

PCR Primers

FORWARD: T3 20mer

BACKWARD: T7 21mer (backward)

Plate: 068 row: E column: 05

Seq primer: T3 20mer

High quality sequence stop: 279.

## FEATURES

source

1..279  
/organism="Vitis vinifera"  
/mol\_type="mRNA"  
/db\_xref="taxon:29760"  
/clone="VVD068E05"  
/tissue\_type="berries"  
/dev\_stage="mixed; 8, 9, 11, 13, 15, 16 weeks daf"  
/clone\_lib="An expressed sequence tag database for abiotic  
stressed berries of Vitis vinifera var. Chardonnay"  
/note="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site 1:  
EcoRI; Site 2: XhoI"

## ORIGIN

Alignment Scores:  
Pred. No.: 6.14e+03 Length: 279  
Score: 27.00 Matches: 6

Percent Similarity: 100.00%  
 Best Local Similarity: 100.00%  
 Query Match: 100.00%  
 DB: 6  
 Gaps: 0

US-10-014-101B-41 (1-6) x CB921222 (1-279)

QY 1 lleThrArgAlaArgile 6  
 Db 95 ATTACAGGCGCATGATA 78

RESULT 36  
 CD152034  
 LOCUS  
 DEFINITION  
 ML1-0027T-D047-H07-U.G mRNA linear EST 14-SEP-2003  
 ML1-0027T-D047-H07.G, mRNA sequence.  
 ACCESSION  
 CD152034  
 VERSION  
 CD152034.1 GI:34688914  
 KEYWORDS  
 EST.  
 SOURCE  
 Schistosoma mansoni  
 ORGANISM  
 Schistosoma mansoni  
 Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;  
 Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.  
 REFERENCE  
 1 (bases 1 to 285)  
 Verjovski-Almeida, S., DeMarco, R., Martins, E.A.L., Guimaraes, P.E.M.,  
 Ojopi, E.P.B., Paquola, A.C.M., Piazza, J.P., Nishiyama, M.Y. Jr.,  
 Kitajima, J.P., Adamson, R.E., Ashton, P.D., Bonaldo, M.P.,  
 Coulson, P.S., Dillon, G.P., Farias, L.P., Gregorio, S.P., Ho, P.L.,  
 Leite, R.A., Malaquias, L.C.C., Marques, R.C.P., Miyasato, P.A.,  
 Nascimento, A.L.T.O., Ohlweiler, F.P., Reis, E.M., Ribeiro, M.A.,  
 Sa, R.G., Stukart, G.C., Soares, M.B., Gargioni, C., Kawano, T.,  
 Rodrigues, V., Madeira, A.M.B.N., Wilson, R.A., Menck, C.F.M.,  
 Secubal, J.C., Leite, L.C.C. and Dias-Neto, E.  
 Transcriptional analysis of the acelomate human parasite Schistosoma  
 mansoni  
 Nat. Genet. 35 (2), 148-157 (2003)  
 22879926  
 12973350

COMMENT  
 Contact: Dr. Sergio Verjovski-Almeida  
 Departamento de Bioquímica  
 Instituto de Química - Universidade de São Paulo  
 Av. Prof. Lineu Prestes 748 sala 1200, 05508-900 São Paulo - SP,  
 Brasil  
 Tel: +55-11-3091-2173  
 Fax: +55-11-3091-2186  
 Email: verj@iq.usp.br  
 This sequence was derived from the FAPESP Schistosoma mansoni EST  
 Genome Project. All sequences in the project were assembled and  
 annotated. This entry and all the assembled sequences can be seen  
 in the following URL: <http://bioinfo.iq.usp.br/schisto/>  
 Plate: ML1-0027T-D047 row: 7 column: H.

FEATURES  
 Location/Qualifiers  
 1..285  
 /organism="Schistosoma mansoni"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:6183"  
 /clone="ML1-0027T-D047-H07.G"  
 /sex="mixed pool"  
 /dev\_stage="miracidium"  
 /clone\_lib="ML1-0027"  
 /note="Vector: TopoBlunt"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 6.27e+03 Length: 285  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00%  
 Best Local Similarity: 100.00%  
 Query Match: 100.00%  
 DB: 6  
 Gaps: 0

US-10-014-101B-41 (1-6) x CD152034 (1-285)

QY 1 lleThrArgAlaArgile 6  
 Db 29 ATAACAGGCGCATGATA 46

RESULT 37  
 AW870159  
 LOCUS  
 DEFINITION  
 NXNV\_124\_G12\_F Nsf Xylem Normal wood Vertical Pinus taeda CDNA  
 clone NXNV\_124\_G12 5', mRNA sequence.  
 ACCESSION  
 AW870159  
 VERSION  
 AW870159.1 GI:8004212  
 KEYWORDS  
 EST.  
 SOURCE  
 Pinus taeda (loblolly pine)  
 ORGANISM  
 Pinus taeda  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.

REFERENCE  
 1 (bases 1 to 286)  
 Sederoff, R.  
 TITLE  
 Molecular Basis of Wood Formation in the Pine Megagenome  
 JOURNAL  
 Unpublished (2000)  
 COMMENT  
 Contact: Sederoff, Ron  
 Forest Biotechnology  
 North Carolina State University  
 840 Main Campus Drive, Centennial Campus, Campus Box 7247, Raleigh,  
 NC 27695, USA  
 Tel: 919 515 7800  
 Fax: 919 515 7801  
 Email: ron.sederoff@ncsu.edu, jerri.johnson@ncsu.edu  
 Please see <http://web.ahc.umn.edu/biodata/nsfpine/> for further  
 information.  
 Seq primer: T3.

FEATURES  
 Location/Qualifiers  
 1..286  
 /organism="Pinus taeda"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3352"  
 /clone="NXNV\_124\_G12"  
 /clone\_lib="Nsf Xylem Normal wood Vertical"  
 /note="Vector: BlueScript SK; Site 1: Eco RI; The  
 sequences contain a CDNA adapter between the EcoRI site  
 and the start of the EST. The adapter sequence is  
 'AATTCGCGCAGAG'."

## ORIGIN

Alignment Scores:  
 Pred. No.: 6.29e+03 Length: 286  
 Score: 27.00 Matches: 6  
 Percent Similarity: 100.00%  
 Best Local Similarity: 100.00%  
 Query Match: 100.00%  
 DB: 2  
 Gaps: 0

US-10-014-101B-41 (1-6) x AW870159 (1-286)

QY 1 lleThrArgAlaArgile 6  
 Db 87 ATCACTAGGCGCAGATC 104

RESULT 38  
 BB336254  
 LOCUS  
 DEFINITION  
 BB336254 RIKEN full-length enriched, 10 days neonate medulla  
 oblongata Mus musculus CDNA clone BB30048F12 3', mRNA sequence.  
 ACCESSION  
 BB336254  
 VERSION  
 BB336254.1 GI:9045017  
 KEYWORDS  
 EST.  
 SOURCE  
 Mus musculus (house mouse)  
 ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
 1 (bases 1 to 287)  
 Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T.,  
 Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,



RESULT 40  
BM348722/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

BM348722 289 bp mRNA linear EST 15-JAN-2002  
MEST299-G08.T3 ISUMS-RN Zea mays cDNA clone MEST299-G08 3', mRNA  
sequence.  
BM348722  
BM348722.1 GI:18173334  
EST.  
Zea mays  
Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 289)  
Wen,T.J., Qu,F., Guo,L., Ashlock,D.A and Schnable,P.S.  
Expressed Sequence Tags from B73 Maize: various stages and tissues  
including seedlings treated with a variety of hormones  
Unpublished (2001)  
Contact: Patrick S. Schnable  
Schnable Laboratory  
Iowa State University  
G405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA  
Tel: 515-294-0975  
Fax: 515-294-2299  
Email: schnable@iastate.edu  
Individual basecall and confidence value were assigned using the  
phred software,  
([http://depts.washington.edu/ventures/collabtr/direct/index.htm#b  
rt](http://depts.washington.edu/ventures/collabtr/direct/index.htm#b<br/>rt)). Overall sequence quality assessment and vector trimming were  
conducted using the Lucy software (<http://www.tigr.org/softlab/lucy>).  
Lucy parameters were set to ensure an overall trimmed quality of  
97.5% or better without any vector fragments in the chosen  
high-quality region of each sequence. Low-quality bases between the  
poly-T and the high-quality region were replaced with N's to serve  
as spacers.  
PCR Primers  
FORWARD: primer T7-1 (AA TAC GAC TCA CTA TAG)  
BACKWARD: primer T3 (ATT AAC CCT CAC TAA AG)  
Seq primer: primer T3 (ATT AAC CCT CAC TAA AG).

Location/Qualifiers  
1. .289  
/organism="Zea mays"  
/mol\_type="mRNA"  
/cultivar="B73"  
/db\_xref="taxon:4577"  
/clone="MEST299-G08"  
/tissue\_type="mixed"  
/lab\_host="DH10B"  
/clone\_lib="ISUMS-RN"  
/note="Vector: pT73PAC; Site 1: EcoRI; Site 2: NotI;  
Tissues: Germinated seed and seedlings (1, 2, 8, 11 DAG),  
Mixed mature tissues (17, 21, 38, 69, 77 DAG), Kernels  
(3, 5, 10, 15, 20, 25, 30, DAP), Adventitious roots (65  
DAG), Tassel (3-39 cm, 53 and 56 DAG), Immature ear  
(0.2-3.0 cm, 53, 56, 59 DAG), Husk (73 DAG), Silk,  
unpollinated first ear, ear shank, etiolated seedlings,  
callus, Cycloheximide-treated callus, Anaerobic treated  
seedlings, NAA (a-Naphthalene acetic acid)-treated  
seedlings, Kinetin-treated seedlings, ACP  
(1-aminocyclopropane-1-carboxylic acid)-treated seedlings,  
Brassinolide-treated seedlings, ABA (Abscissic  
acid)-treated seedlings, GA (Gibberellic acid)-treated  
seedlings, JA (Jasmonic acid)-treated seedlings. ds-cDNA  
molecules were generated as follows. First-strand cDNA was  
prepared from oligo-dT selected mRNA by priming with a  
NotI oligo-dT primer (5'  
AATCGAGAAATTCGGCCGACGAGAAATTTTTTTTTTTTTTTT). The  
resulting DNA:RNA hybrid was treated with RNase H and used  
as a template for DNA PolI-catalyzed second strand  
synthesis. After the addition of EcoRI adaptors, the  
ds-cDNAs were digested with NotI and size-selected. The  
resulting molecules were directionally cloned into the

ORIGIN  
Alignment Scores:  
Pred. No.: 6.36e+03 Length: 289  
Score: 27.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0  
US-10-014-101B-41 (1-6) x BM348722 (1-289)  
QY 1 IleThrArgAlaArgile 6  
Db 136 ATTACACGTGCACGCAATT 119  
Search completed: February 18, 2005, 07:01:04  
Job time : 713.52 secs